

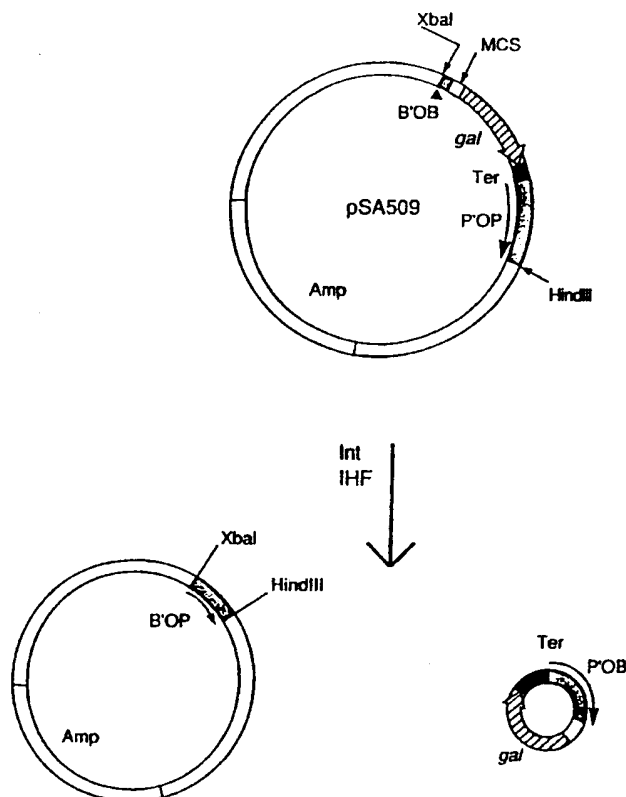


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> <b>C12N 15/10, 15/70, 1/21</b> <b>// (C12N 1/21, C12R 1:19)</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 94/09127</b> <b>(43) International Publication Date:</b> 28 April 1994 (28.04.94)
<b>(21) International Application Number:</b> PCT/US93/09918 <b>(22) International Filing Date:</b> 15 October 1993 (15.10.93) <b>(30) Priority data:</b> 07/963,328                      16 October 1992 (16.10.92)      US <b>(71) Applicant:</b> THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Box OTT, Bethesda, MD 20892 (US). <b>(72) Inventors:</b> ADHYA, Sankar, L. ; 14400 Kings Grant Street, Gaithersburg, MD 20878 (US). CHOY, Hyon, E. ; 3 Pookshille, Bethesda, MD 20814 (US).		<b>(74) Agents:</b> FEILER, William, S. et al.; Morgan & Finnegan, 345 Park Avenue, New York, NY 10154 (US). <b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>

**(54) Title:** SUPERCOILED MINICIRCLE DNA AS A UNITARY PROMOTER VECTOR**(57) Abstract**

A system to construct supercoiled "minicircle" DNA containing a single promoter by the use of the integrative recombination system of bacteriophage lambda is described. The supercoiled DNA minicircles are templates which carry a single promoter. Such minicircles serve as templates to study the steps of transcription initiation from a unitary promoter and to synthesize both full length transcripts and aborted transcripts under physiological conditions in the absence of other contaminating promoter activities. The invention is directed to *in vitro* transcription of promoters, such as the *E. coli gal* promoter, using such minicircle DNA templates. The full-length transcripts from *gal* P1 and P2 promoters responded to cAMP-CRP in a manner identical to that observed *in vivo*. In addition, an *in vitro* transcription assay, using minicircle DNA containing the *gal* promoter with *lac* operators, was employed to elucidate the molecular mechanism of repression. Lac repressors, which associate into a tetramer and form a DNA loop, repressed transcription from P1 and P2, while a non-looping Lac repressor mutant failed to show normal repression of both of the *gal* promoters.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic	KR	Republic of Korea	SD	Sudan
CG	Congo	KZ	Kazakhstan	SE	Sweden
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovak Republic
CM	Cameroon	LU	Luxembourg	SN	Senegal
CN	China	LV	Latvia	TD	Chad
CS	Czechoslovakia	MC	Monaco	TC	Togo
CZ	Czech Republic	MG	Madagascar	UA	Ukraine
DE	Germany	ML	Mali	US	United States of America
DK	Denmark	MN	Mongolia	UZ	Uzbekistan
ES	Spain			VN	Viet Nam
FI	Finland				

## SUPERCOILED MINICIRCLE DNA AS A UNITARY PROMOTER VECTOR

### FIELD OF THE INVENTION

5           The present invention relates to the production of supercoiled DNA minicircles which can be employed *in vitro* as a purified system to elucidate, for example, the various aspects of transcription, namely, initiation, elongation, and termination. The present system closely resembles the molecular events which  
10          occur *in vivo* and provides supercoiled DNA plasmid circles which carry only one promoter and which are considered as unitary promoter vectors.

### BACKGROUND OF THE INVENTION

15           Transcription involves the synthesis of RNA complementary to one strand of a DNA duplex, i.e. the "template" DNA, by the enzyme RNA polymerase. Transcription can be divided into three stages: (i) initiation, (ii) elongation and (iii) termination.

20           Initiation of transcription begins with the binding of RNA polymerase to the double-stranded DNA. During the initiation of transcription, RNA polymerase goes through at least four stages: (i) initial binding to the promoter (closed complex formation), (ii) isomerization of the closed complex to make the template strand available for base pairing with ribonucleotides (open  
25          complex formation), (iii) idling during which RNA oligomers are made and aborted by this initial transcribing complex and (iv) clearance of the promoter concomitant with RNA chain elongation (1-7).

30           During initiation of transcription, when RNA polymerase makes aborted oligomers before clearing the promoter as an elongating complex, the polymerase returns to the isomerized complex and recycles the initiation reactions. During these steps, RNA polymerase undergoes a series of conformational changes that may be subject to regulation (7, 8). *In vitro* transcription studies have shown  
35          the synthesis of considerable amounts of aborted transcripts of less than ten

- 2 -

nucleotides in length by many promoters. These results have underscored the importance of the idling and promoter clearance steps, in addition to closed and open complex formation, in the regulation of transcription initiation (6, 7).

The measurement of closed complex formation and subsequent open complex formation routinely uses an abortive initiation assay (9). A regulatory assessment of the idling and the promoter clearance steps requires a transcription assay that employs a DNA template with a single promoter (i.e. a unitary promoter DNA template) so that both aborted and full length transcripts can be analyzed directly and simultaneously without interference by the products of other promoters. Since topological conformation of DNA can greatly influence the interactions between DNA and the transcriptional regulatory proteins (10-12, 41, 42) in that the efficiency of promoters is usually influenced by the degree of supercoiling, it would be relevant if the unitary promoter DNA template is supercoiled.

Communication between proteins bound to spatially separated DNA control elements for the purpose of controlling gene transcription by looping out the intervening DNA segment has been shown or implicated for both prokaryotic and eukaryotic systems (reviewed in 34). DNA looping participates in repression as well as in activation of transcription initiation. An activator or a repressor can potentially modulate transcription from a promoter at the level of (i) RNA polymerase (or an activator protein) binding, (ii) RNA polymerase isomerization, (iii) formation of the initial phosphodiester bonds, or (iv) promoter clearance by RNA polymerase (7, 8).

In the *gal* operon of *Escherichia coli*, two independent promoters, *P*<sub>1</sub> and *P*<sub>2</sub>, which are located 5 base-pairs apart, drive the expression of that operon (16-18). cAMP and its receptor protein (CRP) complex modulate the two promoters in opposite directions: cAMP•CRP complex stimulates the *P*<sub>1</sub> promoter but represses the *P*<sub>2</sub> promoter. Repression of transcription from the two partially overlapping promoters, *P*<sub>1</sub> and *P*<sub>2</sub>, requires DNA looping generated by interaction of Gal repressor molecules bound to two operators, *O*<sub>E</sub> and *O*<sub>I</sub>, which encompass the promoters and are separated by 11 B-DNA helical turns (23, 25,

- 3 -

38). DNase competition protection experiments in the *gal* system have shown that Gal repressor binding to  $O_E$  and  $O_I$  does not prevent binding of RNA polymerase or of cyclic AMP-cyclic AMP receptor protein complex (CRP) which is an activator of the *P1* promoter (39, 40). These results indicate that Gal repressor acts at one of the latter three steps described above.

5

It is plausible mechanistically that a DNA loop structure "locks" the RNA polymerase at the promoter, thereby establishing repression without any direct contact between repressor and RNA polymerase. Alternatively, repression can be established through communication(s) between repressor and allosteric site(s) in the RNA polymerase as part of a nucleoprotein complex (8). Such contacts in the complex make RNA polymerase assume an idle form in repression. DNA looping juxtaposes the necessary components in a geometrically proper complex.

10

15

The study of the mechanism of repression of the *lac* promoter by  $LacI^+$  repressor in purified systems has led to conflicting conclusions. Although it was originally believed that Lac repressor acts at the first step by inhibiting closed complex formation in the *lac* operon (47), it has been suggested that with wild type *lac* promoter, the repressor blocks a step after the formation of a closed promoter complex (48). More specifically, from studying the cAMP independent *lacUV5* promoter at high concentrations of Lac repressor and RNA polymerase, Straney and Crothers (49) suggested that Lac repressor blocks the second step, the isomerization of the RNA polymerase-*lac* promoter complex to an open form. In contrast, Lee and Goldfarb (50) have concluded that in the *lacUV5* promoter, Lac repressor creates a kinetic barrier in the fourth step, the promoter clearance step, which leads to an idling and aborted transcript-producing RNA polymerase.

20

25

Elongation occurs when RNA polymerase clears the promoter, moves along the DNA, unwinds the DNA duplex to expose a new segment of the template DNA in single-stranded condition, and extends the growing RNA chain, thus forming an RNA-DNA hybrid in the unwound region, producing an elongation ternary complex. As the RNA polymerase moves, the RNA that was previously made is displaced from the template DNA and the DNA duplex reforms in that

30

35

- 4 -

region.

During termination of transcription, the last base is added to the RNA chain, the RNA and RNA polymerase are released from the template DNA, and the DNA duplex reforms.

5

## SUMMARY OF THE INVENTION

The present invention provides a DNA construct, i.e. a plasmid, which is designed to contain a selectable marker, such as ampicillin resistance (Amp); the bacterial and lambda attachment sites *attB* and *attP* (i.e. B'OB and P'OP); a multiple cloning site (*mcs*); a DNA sequence or segment of interest  
10 cloned into the *mcs*, such as the promoter and transcription start site of the *gal* gene; and a transcriptional terminator sequence (Ter), for example, the Ter of the *rpoBC* operon of *Escherichia coli*. A linear depiction of the relevant segment of the plasmid construct is as follows: - Amp - B'OB - *mcs* - *gal* - Ter - P'OP - .  
15 The plasmid vector of the invention is introduced into a strain of bacteria which provides the lambda integrase (Int) protein and integration host factors (IHF). The integrase gene, which encodes the Int protein, is regulated by a temperature sensitive lambda repressor, *cI857*. As a result, no Int is produced at the  
20 temperature of 32°C; however, at the temperature of 42°C, the *cI857* repressor is denatured, allowing the expression of the integrase gene and the production of the Int protein. In the present invention, host bacteria containing the plasmid vector of the invention are grown at 32°C. When the temperature is elevated to 42°C, the Int  
25 protein is expressed and a supercoiled DNA minicircle is created by the excision of the OB - *mcs* - *gal* - Ter - P' segment of the plasmid and is circularized by ligation of the two ends into an isolatable miniplasmid or "minicircle". The Int protein and IHF catalyze the entire process by binding to B'OB and P'OP. Since  
30 the Int protein is somewhat heat labile, the temperature of the bacterial culture is returned to 32°C from 42°C for optimal yield.

35

The present invention serves as a means for the analysis of the precise and specific transcriptional properties of promoter and/or terminator present in the resulting minicircle DNA. The present invention also provides a method of

- 5 -

- *in vitro* transcription using purified bacterial RNA polymerase, other transcription factors, and the DNA minicircle as a template.

The present invention is also applicable to the study of eukaryotic including mammalian, yeast and plant, transcription systems. For eukaryotic systems, eukaryotic promoters, initiators, terminators and/or other transcriptional elements are used in the *mcs* of the plasmid vector and the appropriate eukaryotic enzymes and/or regulatory factors are supplied in the transcription reaction mixture.

The present invention is also applicable to the study of hybrid, homologous or heterologous, prokaryotic and eukaryotic transcription systems. For hybrid systems, combinations of eukaryotic or prokaryotic promoters, initiators, terminators and/or other transcriptional elements are used in the *mcs* of the plasmid vector and the appropriate eukaryotic and prokaryotic enzymes and/or regulatory factors are supplied in the transcription reaction mixture.

The invention generates supercoiled DNA miniplasmids, also called minicircles, which comprise only one promoter, or, as in the case of the *gal* operon, two overlapping promoters. The DNA minicircles are frequently more efficient substrates for *in vitro* transcription than are linear DNA segments. A supercoiled miniplasmid with unique promoter(s) can avoid the use of oligonucleotides as primers which are required to analyze promoters from both linear and circular DNA templates which contain multiple promoters.

It is an objective of the present invention to provide small supercoiled DNA minicircles which are easy to make and isolate in high yields. Such supercoiled DNA minicircles are more like naturally-occurring DNA, as opposed to linear DNA fragments as templates for RNA transcription.

It is a another object of the invention to provide supercoiled DNA minicircles which carry only the promoter of interest so as to alleviate any interference from other, extraneous promoters, or, for example, from an origin of replication having promoter activities, that may be present in other plasmids.

It is still another objective of the invention to generate a purified system in which to produce small DNA minicircles for the identification or analysis

- 6 -

- of the function of particular genes or portions of genes, preferably promoters, without competition from other genes or from other promoters, or promoter-like sequences, which may vie for the binding of RNA polymerase and which may obscure the analyses.

Another objective is to develop a DNA vector which would allow the simple detection and assay of the entire array of transcription products -- both complete RNA, as well as aborted RNA oligomers, produced either by single cycle or by reiterative cycle reactions -- in a purified system with supercoiled DNA. The supercoiled DNA minicircles carrying only one promoter followed by a terminator can yield only one full-sized RNA of discrete size for easy identification.

A further objective of the invention to provide an *in vitro* unitary transcription system which does not have to use oligonucleotide primers to study transcription initiation.

It is yet another objective of the invention to generate a purified system in which to produce small DNA minicircles useful as substrates for the identification and/or analyses of DNA or RNA sequences, proteins and/or other components involved in the coupling of transcription and translation. When the small DNA minicircles of the invention are used in such a manner, the appropriate enzymes and/or regulatory factors are supplied in an appropriate transcription/translation mixture.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1:** Arrangement of the functional elements of the unitary promoter plasmid vector, pSA508. The symbols and abbreviations as shown on the vector construct are as follows: *XbaI*: unique restriction site for endonuclease *XbaI*; B'OB: bacterial attachment site attB'OB; *mcs*: multiple restriction endonuclease sites for cloning; Ter: transcription termination signal; P'OP: lambda phage attachment site attP'OP; *HindIII*: unique restriction site for endonuclease *HindIII*; and Amp: gene encoding ampicillin resistance.

**Figure 2:** DNA profiles of plasmid pSA508 before and after Int/IHF-mediated *in*



- 7 -

*in vivo* recombination displayed on a 1.0% agarose gel. Lane 1 shows the 1 kb ladder (BRL). Lane 2, plasmid DNA before recombination. Lanes 3 and 4, DNA profile of the recombination products. Lane 4 was loaded with five fold more DNA sample than lane 3 to make the minicircle product "a" clearly visible. Lane 5, the monomer DNA minicircle after purification. Arrows: a, monomer minicircle; b, the larger product circle; c, a circle with two units of larger product circle; S1, monomer substrate; and S2, dimer substrate.

**Figure 3:** *gal* promoter of *E. coli* in the minicircle generated from pSA509. O<sub>E</sub> and O<sub>I</sub> represent external and internal operator, respectively; Ter is the transcription terminator.

**Figure 4:** Effect of cAMP concentrations on transcription using the vector minicircle (from pSA508) or *gal* minicircle (from pSA509) as template. CRP was present at 100 nM. Transcription products were analyzed on an 8% (A) and on a 25% (B) urea-polyacrylamide gel. Lanes 1-5 in A and 1-5 in B were loaded with RNA made from vector minicircle; lanes 6-10 in A and 6-10 in B contained RNA made from *gal* promoter containing minicircle. The full-length transcripts on the 8% gel were quantified using a beta-scanner, and the total counts per minute (cpm) of each band were plotted against cAMP concentrations (C). The open circles represent full-length RNA made from P1; the closed circles represent full-length RNA made from P2.

**Figure 5:** Rate of *gal* RNA synthesis. A reaction mixture containing *gal* promoter DNA template and RNA polymerase was preincubated at 37°C for 5 min after which heparin was added to 100 µg/ml. Transcription reactions were initiated by the addition of nucleoside triphosphate substrates 1 minute after heparin challenge and were terminated at different times as indicated. The reaction products were analyzed on 8% (A) and 25% (B) urea-polyacrylamide gels for full-length and aborted transcripts, respectively.

**Figure 6:** *gal* transcription from minicircle DNA templates carrying the wild type *gal* promoter, pSA509, (lanes 1-5) and mutant *gal* promoter, pSA515, (lanes 6-10). The reaction conditions are the same as in Figure 3. The top panel (A) shows the full-length transcripts on an 8% gel urea-polyacrylamide and the bottom panel (B)

- 8 -

shows aborted transcripts on a 25% urea-polyacrylamide gel.

**Figure 7:** Lac repressor ( $\text{LacI}^+$ ) was titrated *in vitro* in the presence of absence or 0.2 mM cAMP. CRP was present at 100 nM in all lanes. The transcription products were displayed on an 8% (panel A) and a 25% (panel B) denaturing urea-polyacrylamide gel. The control reactions were carried out in the presence of 1 mM IPTG. The full length transcripts (panel A) were quantified using a beta-scanner, and the fraction of RNA in each lane compared to RNA made in the presence of IPTG, was plotted as a function of Lac repressor concentration (panel C). Open circles represent *P1* activity and closed circles represent *P2* activity.

**Figure 8:** The mutant Lac repressor ( $\text{LacI}^{\text{di}}$ ) was titrated *in vitro* in the presence or absence of 0.2 mM cAMP. CRP was present at 100 nM. The transcription products displayed on an 8% (panel A) and a 25% (panel B) denaturing urea-polyacrylamide gel. The control reactions were carried out in the presence of 1 mM IPTG. The full length transcripts (panel A) were quantified using a beta-scanner, and the fraction of RNA in each lane, compared to RNA made in the presence of IPTG, was plotted as a function of Lac repressor concentration (panel C). Open circles represent *P1* activity and closed circles represent *P2* activity.

**Figure 9:** Gal repressor ( $\text{GalR}$ ) was titrated *in vitro* in the presence or absence of 0.2 mM cAMP. CRP was present at 100 nM. The full length transcripts were quantified using a beta-scanner, and the amount of RNA relative to that made in the presence of the inducer (10 mM D-galactose) was plotted as a function of Gal repressor. The DNA templates used in this study carried *gal* promoter with wild type *gal* operators (pSA509, panel A); with wild type external operator and constitutive internal operator (pSA511, panel B); with constitutive external operator and wild type internal operator (pSA512, panel C). Open circles represent *P1* activity and closed circles represent *P2* activity.

30

35

- 9 -

- ° **Figure 10:** Arrangement of the functional elements of the unitary promoter plasmid, pSA509, and the generation of a DNA minicircle as a result of Int/IHF-mediated recombination. The symbols, abbreviations, and the details of the recombination reaction are as follows: *XbaI*: unique restriction site for endonuclease *XbaI*; B'OB: bacterial attachment site attB'OB; *mcs*: multiple  
5 restriction endonuclease sites for cloning; *gal*: a DNA segment carrying the *gal* promoter and operator loci; Ter: transcription termination signal; P'OP: lambda phage attachment site attP'OP; *HindIII*: unique restriction site for endonuclease *HindIII*; and Amp: gene encoding ampicillin resistance.

10

### DESCRIPTION OF THE INVENTION

One aspect of the invention is the development of a plasmid vector which generates supercoiled unitary promoter DNA templates, called "minicircles",  
15 convenient for *in vitro* transcription study of individual promoters under physiological conditions. The minicircles are generated *in vivo* by taking advantage of the mechanism of site-specific integrative recombination of bacteriophage lambda (13-15). These minicircles are supercoiled, making them  
20 natural substrates for transcription from the cloned promoter. Since these unitary promoter templates do not carry any other transcription initiation sites, all aborted and full length transcripts originate from the cloned promoter, which allows direct qualitative and quantitative analysis of the transcripts by gel electrophoresis. Thus, these minicircles may be used to study various aspects of transcription, including  
25 aspects of transcription initiation, e.g. idling and promoter clearance, as exemplified herein.

Examples of vectors which can be useful in creating source plasmids to generate supercoiled unitary promoter DNA templates of the invention include  
30 those which can be modified to contain the - B'OB - *mcs* - P'OP - sequences and which have the following properties: a selectable antibiotic resistance marker and multicopy in nature, e.g. pIBI24, which is exemplified herein, as well as vectors such as pBR322, and the like.

35

Less preferably, the supercoiled unitary promoter DNA templates of

- 10 -

the invention can also be made *in vitro* by one of the following two ways: (i) by carrying out the site-specific integrative recombination reaction *in vitro* with purified Int and IHF proteins, and (ii) by purifying a small DNA fragment carrying the promoter-terminator segment, and then ligating the two ends using DNA ligase *in vitro*. However, these processes are laborious and are not cost-effective.

Examples of suitable host cells include those which contain and express, or which can be modified to contain and express, lambda Int and integration host factor (IHF) proteins, e.g. *E. coli* strain SA1751 which has been transformed to contain the plasmid pSA50X, a plasmid containing the attB'OB - attP'OP segment, which is exemplified herein, and other similarly constructed cells.

Examples of promoters which may be cloned and studied in the present invention include regulatable or non-regulatable promoters, including, but not limited to, bacterial promoters, such as, for example, *lac*, *lacUV5*, *lpp*, *mal*, and *trp*; bacteriophage promoters, such as, for example, *lambda P<sub>R</sub>*, *lambda P<sub>L</sub>*,  $\phi T7$ ,  $\phi T4$ , and  $\phi SP6$  promoters; and eukaryotic promoters, such as, for example, yeast GAL1 promoter, SV40 early or late promoter, the mouse mammary tumor virus (MMTV) promoter, and the metallothionein (MT) promoter.

Examples of terminators which may be cloned and used in the present invention include, but are not limited to, factor-independent or factor-dependent bacterial terminators such as, for example, *rpoBC* (rho-independent); bacteriophage terminators, such as, for example, *lambda t<sub>L</sub>*, *lambda t<sub>R</sub>*; and eukaryotic terminators, such as, for example, yeast SUP4 terminator. Factor-independent terminators are particularly useful in the present invention. In general, for non-hybrid constructs, prokaryotic terminators are used in conjunction with prokaryotic promoters and eukaryotic terminators are used in conjunction with eukaryotic promoters.

Examples of regulatory elements that may be cloned and studied in the present invention include, but are not limited to, bacterial operators, such as, for example, *gal*, *lac*, *trp*, *aroH*, and *trpR*; bacterial activator sequences, such as, for example, the CRP binding site and OxyR binding site; and eukaryotic response

- 11 -

- elements and enhancers, such as, for example, glucocorticoid response element (GRE), estrogen response element (ERE), metallothionein response element (MRE), heat shock response element (HSRE), the SV40 enhancer (72 bp repeat), the immunoglobulin enhancer, and yeast upstream activator sequences (UAS).

5 When any of the above mentioned promoters, terminators or regulatory elements are factor-dependent, the factors are included in the reaction mixture. Alternatively, the reaction mixture may be used as a means of identifying such factors.

10 Furthermore, any of the above mentioned gene segments, i.e., promoters, terminators and regulatory elements, may be cloned into the construct of the invention in a form similar to their naturally-occurring counterparts, or they may be cloned into the construct of the invention in hybrid form. Examples of such hybrid segments which may be cloned include, for example, hybrid bacterial  
15 promoters from different bacterial genes, e.g. the *tac* promoter (a combination of the *trp* and *lac* promoters); hybrid bacterial promoter and regulatory sequences, e.g. the *gal* promoter and the *lac* operator (exemplified herein); hybrid bacteriophage and bacterial gene segments, e.g. a T7 bacteriophage promoter and the *lac* operator; hybrid eukaryotic (or viral) and bacterial gene segments, e.g. a  
20 SV40 promoter and the *lac* operator; and hybrid bacteriophage and eukaryotic (or viral) gene segments, e.g. a T7 bacteriophage promoter and the yeast *gal4* operator or a SV40 promoter or yeast promoter and a bacteriophage polymerase coding region. Similarly, various terminators may be used with various other segments in  
25 the constructs of the invention.

The present invention may also be used to study the involvement of DNA looping in repressor action and to determine the precise stage of repressor action *in vitro*, as exemplified herein.

30 To elucidate the mechanism underlying the repression of transcription initiation from the overlapping *gal* promoters, *P1* and *P2* of the *gal* operon of *E.coli*, supercoiled DNA minicircles which contained only the *gal* promoter segment and adjoining DNA control elements, followed by a bacterial terminator of the *rpoC* gene, were produced and employed using the methods and  
35

- 12 -

plasmids of the present invention (e.g. Figure 3). These DNA minicircles, containing only *gal* and no other promoters, greatly facilitated investigating the mechanism of both activator (CRP) and Gal repressor action on the synthesis of both aborted and full length transcripts from the *gal* promoters in the same assay. These minicircles were used to study idling and promoter clearance in the *gal* operon of *Escherichia coli*.

Interestingly, the *gal* operon can be repressed not only by Gal repressor, but also by Lac repressor, provided that both bipartite *gal* operators are substituted by *lac* operators (37). Since topologically superhelical conformations of DNA have been shown to enhance the interactions between Lac repressor and other transcriptional regulatory proteins to their cognate DNA sites in the formation of DNA loops (10-12, 41, 42), the effect of repressors on *gal* transcription was studied using the supercoiled "minicircle" DNA templates of the present invention. The mechanism of repression of *P1* and *P2* were investigated simultaneously using both the Gal and Lac repressors.

In this system, the complete transcripts from *P1*, an activator dependent promoter, and *P2*, a factor-independent promoter, were synthesized at about equal efficiencies in the absence of cAMP•CRP, in agreement with the results obtained *in vivo* (19) (see Figure 4). The presence of cAMP•CRP inversely regulated the activities of *P1* and *P2*. cAMP activated *P1* and inhibited *P2* transcription (see Figure 4), also in agreement with the results obtained *in vivo*, thus making the system physiological and suitable for further regulatory studies.

In this system, in addition to the synthesis of the complete transcripts, excess amounts of transcription aborted after making trimers to hexamers (16). A large majority of the aborted oligomers initiated mainly at *P2*. Thus, this system allowed the determination of the stage at which repressor brings about inhibition of RNA synthesis in the *gal* system, as discussed below.

A requirement of DNA looping mediated by repressor bound to the spatially separated operator elements,  $O_E$  and  $O_I$ , for repression of both *P1* and *P2* has been demonstrated *in vivo* (37, 38). The use of the purified system of the invention revealed that the complete repression requires an interaction of the

- 13 -

° repressors bound to the operators rather than the mere occupation of the operators by the repressors. Wild type  $\text{LacI}^+$  repressor, which shows DNA looping by electron microscopy (38, 46) and gel electrophoresis (45), resulted in normal repression of  $P_1$  and  $P_2$  on a *gal* DNA template containing *lac* operators, whereas  
5  $\text{LacI}^{\text{mi}}$ , a non-looping Lac repressor mutant, showed incomplete repression of  $P_1$  and no repression of  $P_2$ .

Gal repressor has been suggested to inhibit transcription from the *gal* promoters at a post RNA polymerase binding step (8, 39). The use of the purified system of the invention has shown that the wild type Lac repressor established  
10 complete repression of not only full length, but also aborted transcripts (Figure 7B), which implied that the repressor inhibited the first phosphodiester bond formation or a step prior to that, i.e., the formation or the activity of the initial transcribing complex (ITC) (7) in the *gal* system.

15 Interestingly, the repression of *gal* DNA expression by  $\text{GalR}^+$  repressor was incomplete in the purified transcription assay of the present invention. The  $P_1$  promoter activity was repressed 80% or less, while  $P_2$  promoter activity was not inhibited at all (see Figure 9A). This behavior of  $\text{GalR}^+$   
20 resembles remarkably the pattern of repression brought about by the nontetrameric, looping-defective  $\text{Lac}^{\text{mi}}$  mutant (38) (see Figure 8), and suggests that purified  $\text{GalR}^+$  protein suffers the same non-looping property that  $\text{LacI}^{\text{mi}}$  mutant protein possesses. Both the mutant Lac repressor and wild type Gal repressor bind to  
25 operators normally as dimers but fail to associate into a tetramer/operator complex (36, 44, 45, 51). Under the conditions used in the present transcription assay, the corresponding operators are fully occupied by mutant Lac repressor and wild type Gal repressor (45, 51). Since Gal repressor has been shown to repress effectively both  $P_1$  and  $P_2$  in a crude cell extract (S-30) system (52), it appears that an  
30 element or a feature of the cell extract is missing in the purified system. The missing component or condition very likely aids association of dimeric repressor bound to operator to form DNA looping in the purified system. In this regard, Gal repressor is different from Lac repressor which can be tetramerized by itself  
35 through a leucine minizipper (53).  $\text{LacI}^{\text{mi}}$  used in this invention is missing the C-

- 14 -

terminal end containing the leucines. Gal repressor does not appear to have the corresponding leucines. Since the dimeric Lac<sup>adi</sup>, unlike GalR<sup>+</sup>, fails to repress *in vivo*, it appears that the proposed looping-aid for GalR<sup>+</sup> does not help Lac<sup>adi</sup>.

In the present invention, a partial repression of *P1* but not of *P2*, that was observed for GalR<sup>+</sup> or LacI<sup>adi</sup> (see Figure 8 and Figure 9A), was also observed with Gal repressor on the *gal* DNA carrying an intact O<sub>E</sub> and a mutant O<sub>I</sub> (see Figure 9B). Since no repression was observed for a mutant O<sub>E</sub> and intact O<sub>I</sub> template (see Figure 9C), these results clearly showed that the partial repression of *P1* by GalR<sup>+</sup> with intact operators was not the consequence of repressor interaction with both O<sub>E</sub> and O<sub>I</sub>. Rather, the partial repression resulted from the occupation of O<sub>E</sub> alone by a dimeric repressor. This is in agreement with the observations made *in vivo* and in the crude system that an intact O<sub>E</sub> alone allowed a partial repression while intact O<sub>I</sub> alone allowed fully constitutive expression of the operon even in the presence of saturating concentrations of repressor (23, 52). Apparently, this partial repression does not involve the DNA loop.

In the aborted transcription initiation assay, Gal repressor has been shown to cause a 10% decrease in the open complex formation at *P1* with a concomitant and similar amount of increase at *P2* both in the absence and presence of cAMP•CRP (30). Consistent with the results obtained from the present invention, the minor repression of *P1* and activation of *P2* by Gal repressor needs only an intact O<sub>E</sub>. The weak repression of *P1* achieved by Gal repressor or LacI<sup>adi</sup> repressor, as was observed in the absence of DNA looping, may not be through a completely different mechanism. In the framework of the repressor-RNA polymerase contact model described above, it is likely that RNA polymerase bound to *P1* contacts the repressor bound to O<sub>E</sub>, but the repressive contact signal, in the absence of DNA looping, is not strong enough to establish complete repression.

As discussed above, using supercoiled minicircles that carry the *gal* promoter segment, the present invention has demonstrated that (1) the two *gal* promoters have different rate limiting steps of transcription initiation by RNA polymerase; (2) cAMP•CRP switches the *gal* promoter activities in the same quantitative fashion as *in vivo* (19); (3) repression of *gal* operon occurs at a step



- 15 -

- ° prior to first phosphodiester bond formation; (4) repression of *gal* operon *in vitro* requires an interaction between repressors bound to two operators; and (5) while Lac repressor repressed both *P1* and *P2*, repression by Gal repressor, as opposed to the *in vivo* result, is incomplete for *P1* and is totally ineffective for *P2*.

5                   Component(s) or conditions that aid Gal repressor in DNA looping remain to be identified. These component(s) or conditions may be determined by assaying the transcripts obtained after supplementing and/or altering the reaction mixtures containing minicircles of the present invention.

10                   The invention is further described by, but is not intended to be not limited to, the following illustrative examples:

## EXAMPLES

### 15    EXAMPLE 1

#### MATERIALS AND METHODS

##### Construction of plasmids

20                   To make supercoiled minicircle DNA carrying only one promoter *in vivo*, the site-specific integrative recombination mechanism of bacteriophage lambda was used (13, 14, 15). A plasmid vector, called pSA508, was constructed (see Figure 1). This vector contained a DNA segment with multiple cloning sites (*mcs*) and a transcription termination sequence (Ter) between the phage lambda attachment site, *attP'OP* and the corresponding bacterial attachment site, *attB'OB*.  
25                   In the presence of lambda Int and host IHF proteins, pSA508 underwent site-specific recombination resulting in two DNA circles, similar to those as shown for pSA509 in Figure 10. In the case of pSA508, the smaller product (i.e. the minicircle or miniplasmid) was approximately 0.2 kb in size, contained the *mcs* and  
30                   the terminator, and did not carry a promoter.

                  The purpose of using *attB'OB* and *attP'OP* sites in the orientation described herein was to produce a supercoiled miniplasmid containing a single promoter of choice and no other promoter. The arrangement of *attB'OB* and  
35                   *attP'OP* in that particular order produced a DNA minicircle that carried *attBOP'*.

- 16 -

- ° If the arrangement had been *attBOB'* and *attPOP'*, as is normally found in phage lambda (33), the product minicircle would have carried the *attB'OP* site. The "P" component of *attB'OP* carries multiple promoters (27, 28) and would not have been serviceable in the present invention (see Figures 2 and 4).

5 pSA508 was derived from pIBI24 (International Biotechnologies, Inc.) by inserting the following DNA elements between its *EcoRI* and *HindIII* sites in the order as follows: the 46 bp bacterial attachment site *attB'OB* (from -19 to +26, ref. 20), a 49 bp *mcs*, the 54 bp Rho-independent transcription terminator of the *rpoBC* operon of *E. coli*,  
 10 GCAAATAACGTAAAAACCCGCTTCGGCGGGT TTTTATGGGGGGAGTTT  
 AGGG (21) (SEQ ID NO:1), and the 408 bp *attP'OP* originating from pPH54 (from +247 to -160; ref. 22). The 156 bp sequence from the *attB'OB* to the transcription terminator was synthesized using a 380B DNA synthesizer (Applied  
 15 Biosystems). The DNA sequence of the 156 bp DNA fragment is shown below:

20 AATTTCTAGA CCTTCCCGTT TCGCTCAAGT TAGTATAAAA  
 AAGCAGGCTT CAACCGAGCT CGTCGACCCG GGTACCGAGC  
 ATGCATGAAT TCGTCTGCAG CAAATAACGT AAAAACCCGC  
 TTCGGCGGGT TTTTATGG GGGGAGTTTA GGGATC (Seq. ID No:2)

25 Plasmid pSA509 (see Figures 3 and 10) contained a 288 base pair segment of the *gal* promoter (-197 to +91; ref. 23) cloned between the *EcoRI* and *PstI* sites in *mcs* of pSA508. This DNA fragment was obtained by amplifying the corresponding segment of plasmid p291 (23) using primers containing an *EcoRI* site at -197 and a *PstI* site at +91 by polymerase chain reaction. In the presence of lambda Int and host IHF proteins, plasmid pSA509 underwent site-specific recombination resulting in two DNA circles as shown in Figure 10. The smaller  
 30 product (i.e. minicircle or miniplasmid) was approximately 0.4 kb in size and contained the *mcs* and the terminator.

35 Plasmid pSA515 was identical to the pSA509 except that it carried an A:T to T:A transversion at the -11 position relative to the P1 transcription initiation site (16). The minicircles from pSA509 and pSA515 each contained 688

- 17 -

- bp. The recombinant region of each plasmid was confirmed by DNA sequencing using standard methods.

Preparation of minicircle DNA *in vivo* by lambda integrative recombination

5                   Recombinant substrate plasmids, pSA508, pSA509 and pSA515, were introduced into *E. coli* strain SA1751 [*int<sup>+</sup>xis439cI857(cro-chlA)<sub>ΔHI</sub>*] by transformation at 32°C. SA1751 is *E. coli* K-12 strain *Fstr<sup>Δ</sup>lac trp(chlD-pgl)<sub>Δ320</sub>*. In this strain, the *int<sup>+</sup>* gene expression from the cryptic prophage is under the control of temperature sensitive prophage repressor, *cI857*. To prepare DNA  
10 minicircle, SA1751 containing a parental plasmid grown to late log phase at 32°C to accumulate substrate plasmid, was shifted to 42°C for 10 minutes to induce briefly Int synthesis from the cryptic lambda prophage by inactivating the thermolabile repressor, and then cooled to 32°C. After 30 minutes at the lower  
15 temperature, the cells were harvested. The rationale for returning the culture to the lower temperature was to prevent the thermolability of the integrative recombination (24). The ten minute heat shock (42°C) and the thirty minute incubation at 32°C were sufficient for the expression of Int protein for maximal recombination between the *attB'OB* and at *attP'OP* elements under the conditions  
20 of the invention. Total plasmid DNA was isolated by an alkaline lysis protocol (25). The plasmid DNA profile resulting from the recombination showed not only the expected substrate and product plasmids but also various multimerized circles. The monomer minicircles were separated in a 1% agarose gel and extracted by  
25 electroelution in dialysis tubing (25). The extracted DNA was extensively dialyzed against 10 mM Tris acetate, pH 8.0, 0.1 mM EDTA and the DNA concentration was determined spectrophotometrically.

As described above, construction and functional elements of the  
30 plasmids used to generate supercoiled "minicircles" *in vivo* were as follows: The parental plasmid carried a multiple cloning site into which the *gal* promoter segment was inserted. This was followed by a transcription terminator. The promoter-terminator region was located on the plasmid between the lambda phage attachment site, *attP'OP*, and the corresponding bacterial site, *attB'OB*. DNA  
35

- 18 -

minicircles carrying the *gal* promoter followed by the transcription terminator were generated by site-specific recombination between the *attP'OP* and *attB'OB* sites *in vivo* in the host, SA1751, which provided the lambda Int protein and integration host factors, IHF. The minicircles were extracted and purified by gel electrophoresis.

pSA508 was the parental plasmid which contained no promoter at the multiple cloning site (see Figure 1). pSA509 contained a 288 base pair segment of the *gal* promoter (-197 to +91) cloned between the EcoRI and the PstI site of pSA508 (23). pSA510 was identical to pSA509 except that both *gal* operator loci,  $O_E$  and  $O_I$ , were substituted with the consensus *lac* operator sequence: TTGTGAGCGCTCACAA (37) (SEQ ID NO:3). pSA511 and pSA512 were also identical to pSA509 except that the internal operator locus ( $O_I$ ) and the external operator locus ( $O_E$ ) were replaced with the *lac* operator sequence, respectively. In the experiments of Figure 9, an  $O^L$  allele served as an  $O^C$  allele for Gal repressor.

### Proteins

CRP was purified from an *E. coli* strain carrying the *crp*<sup>+</sup> gene on a multicopy plasmid pHA5, using FPLC. RNA polymerase (100% saturated with sigma factor) was purchased from Epicentre Techniques (Madison, WI). Gal repressor (GalR<sup>+</sup>) and wild type (LacI<sup>+</sup>) and mutant Lac repressor (LacI<sup>sd</sup>) were purified as described previously (38, 43, 45).

### Transcription assays

Transcription reactions were carried out as described herein. 2 nM DNA template, 0.1 mM ATP, 0.1 mM GTP, 0.1 mM CTP, 0.01 mM UTP, and 10-20  $\mu$ Ci 5' ( $\alpha$ -<sup>32</sup>P) UTP (1 Ci = 37 GBq) were preincubated in transcription buffer (20 mM Tris acetate, pH 7.8, 10 mM magnesium acetate, 100 mM potassium glutamate) at 37°C for 5 minutes. Transcription was initiated by the addition of RNA polymerase to 20 nM in a total volume of 50  $\mu$ l. To study the effect of cAMP, 100 nM CRP and different amounts of cAMP were included in the preincubation mixture. When variable components such as CRP, repressor,

- 19 -

- cAMP or inducer were present, they were included in the preincubation mix at concentrations described in the Description of the Drawings.

Transcription was terminated after 10 min at 37°C, unless indicated otherwise, by the addition of an equal volume (50 µl) of RNA loading buffer (80% (vol/vol) deionized formamide, 1 X TBE (89 mM Tris borate, pH 8.0, 2 mM EDTA), 0.025% Bromophenol blue, 0.025% xylene cyanole). The mixture was heated at 90°C for 2 minutes and electrophoresed through 8% or 25% polyacrylamide sequencing gels (40 cm x 0.4 mm) containing 8 M urea. RNA transcripts were quantified by determining counts per minutes using a beta-scanner (AMBIS, San Diego, CA).

## EXAMPLE 2

### Generation of supercoiled DNA minicircles

DNA minicircles with or without the *gal* promoters were generated from substrate plasmids by the mechanism of site-specific integrative recombination of phage lambda by integrase (Int) and integration host factor (IHF) as described in Example 1, Materials and Methods. The bacterial *attB'*OB site and the phage *attP'*OP site were placed in the constructs of the invention in a head-to-tail configuration spanning the *mcs* with or without promoter, such that recombination between the two *att* sites would release the intervening 0.4 kb long DNA segment as a circle, called a minicircle, carrying the *attP'*OB site, and a larger circle carrying the *attB'*OP site (Figure 10). The transcription terminator was situated between the *mcs* and the *attP'*OP. Thus, the minicircle carried the *mcs* (with or without a cloned promoter) and the terminator. The two *att* sites in the substrate plasmid were arranged in such a way that the *P* element of *attP'*OP was absent in the minicircle product. The *P* element carried several promoter activities (27,28) that would generate, if present in the minicircles, unwanted RNA products during transcription assays.

Recombination was carried out as described in Example 1, Materials and Methods. The results of the recombination reactions for parental plasmid pSA508 are shown in Figure 2. There was no recombination product before heat

- 20 -

induction of Int synthesis (Figure 2, lane 2). The supercoiled pSA508 was 3.3 kb and migrated as a monomer at a position marked S1 and as a dimer at a position marked S2. S2 was generated by homologous recombination. The Int/IHF-mediated recombination under these conditions was sufficient to convert most of the substrates to products. A longer incubation time did not improve product formation.

The substrates and products of recombination were analyzed by restriction nuclease digestion. Lanes 3 and 4 of Figure 2 show the two recombinant product circles of approximately 2.9 kb (marked "b") and approximately 0.4 kb (minicircle, marked "a"). However, the two lanes showed, in addition, multiple ladders of discrete sizes. The ladder comprising bands of ascending sizes and decreasing intensities, visible between "a" and "b" contained oligomeric forms of the minicircle. Interestingly, another ladder of bands of ascending sizes and decreasing intensities was also observed above the larger product circle "b". The latter ladder represented larger products with increasing numbers of minicircle units. Arrow "c" points to the dimer of the larger product circle, and the ladder above represented the same with increments of the minicircle unit. The ladders were not generated by catenation, but rather by covalent additions of minicircles, as determined by restriction enzyme analysis. The above-described system generated a ladder of supercoiled circles of increasing sizes as determined by gel electrophoresis. Such ladders are useful to create size markers in gel electrophoresis for analysis of supercoiled circles.

The following events resulted in the generation of the ladders: first, Int/IHF-mediated intermolecular *attB'OB x attP'OP* recombinations produced parental oligomeric substrates. Next, such oligomers underwent stepwise intramolecular *attB'OB x attP'OP*, *attP'OP x attP'OB*, or *attP'OB x attP'OB* recombination. The Int-mediated recombination between these three pairs of *att* sites is considerably efficient *in vivo* (26). The employment of a *recA*<sup>-</sup> host for generation of minicircles did not reduce the oligomerization of the products (data not shown). The monomer minicircles, generated *in vivo*, were easily obtained in sufficient quantity for use in transcription assays.

- 21 -

**EXAMPLE 3**cAMP effect on *gal* transcription

Transcription initiated from the *gal* promoters was monitored in the presence of excess CRP (100 nM) and varying concentrations of cAMP using DNA minicircles containing a wild type *gal* promoter segment generated from plasmid pSA509. Minicircles containing no promoter generated from vector plasmid pSA508 were used as control DNA templates. The transcription products were analyzed on 8% and 25% denaturing polyacrylamide gels for visualizing the full length and aborted transcripts, respectively (Figure 4A and B).

The promoter-free vector DNA minicircles were transcriptionally sterile; practically no full or aborted transcripts were seen in 8% or 25% polyacrylamide gels (Figure 4A, lanes 1-5 and 4B, lanes 1-5). The DNA minicircles containing *gal* promoter segment, on the other hand, yielded two distinct full-length transcripts in the absence of cAMP: a 120 nucleotide RNA from the *P1* promoter and a 125 nucleotide RNA from *P2* promoter were seen in the 8% gel (Figure 4A, lane 6). In the presence of increasing concentrations of cAMP, the amount of *P1* transcripts increased, while the amount of *P2* transcripts decreased (lanes 6-10). The *gal*-specific RNA bands constituted more than 90% of the total transcripts. The 54 bp-long stretch of DNA containing the stem and loop structure of *rpoC* terminator was extremely effective in terminating transcription initiated from either *gal* promoter. The total radioactivity of each band was quantified and plotted as a function of cAMP concentration (Figure 4C). As the concentration of cAMP increased, the *P2* product decreased and was nearly undetectable at higher than 20  $\mu$ M cAMP, while the *P1* product increased three-fold. The quantitative effect of cAMP on the synthesis of full length transcripts from the two *gal* promoters was identical to that observed *in vivo* (19).

**EXAMPLE 4**Abortive transcription from the *gal* promoters

Analysis of the transcription products on a 25% polyacrylamide gel showed the synthesis of large amounts of oligomeric transcripts from the *gal*

- 22 -

° promoters described previously (Figure 4B, lanes 6-10, ref. 16). Dimers, trimers, and hexamers were the major products. The intensities of trimers and hexamers, as well as the minor tetramers, were inversely proportional to the concentration of cAMP, suggesting that they originated from *P2*. The pentamer, although made in much smaller amounts, showed a doublet. Its two members responded differently toward cAMP: the upper pentamer decreased while the lower one increased with increasing concentrations of cAMP, suggesting that their origins were from *P2* and *P1*, respectively. The slightly different mobilities of the pentamers from *P1* and *P2* are very likely the result of the differences in their nucleotide compositions. The dimers and heptamers are also believed to be a mixture of transcripts originating from both *P1* and *P2*. The bands seen at the bottom of the gel for both vector and *gal* promoter containing minicircles represented the unincorporated radioactive UTP. Mutant templates were used to confirm the assignment of the respective oligomers to *P2* and *P1* promoters. In a *P1* mutant template, the synthesis of aborted transcripts assigned to *P2* based on cAMP sensitivity remained intact, while that attributed to *P1* disappeared (see below).

The analysis of products of a single-round of transcription was used to determine whether the RNA oligomers were mostly precursors of the full length transcripts or were aborted transcripts. RNA polymerase was pre-incubated for 10 minutes with the template at 37°C in transcription buffer as described above in "Transcription assays". Following a 1 minute heparin (100 µg/ml) challenge, the reaction was initiated with the addition of a nucleoside triphosphate mixture containing radioactive UTP. The reaction was terminated at different times, and the transcripts were analyzed on 8% (Figure 5A) and 25% (Figure 5B) gels. The synthesis of full length transcripts from both *P1* and *P2* reached maximum in about 3 minutes (Figure 5A). Similarly, oligomers originating from both *P1* and *P2* followed the same kinetics as full-length transcripts -- their synthesis increased up to 3 minutes (Figure 5B). The continuous synthesis of aborted and full length transcripts at the same rate in the same relative amounts for 3 minutes in the presence of heparin suggests that transcription initiation by the open complexes at the *gal* promoters is a slow process. The initial transcribing complexes frequently



- 23 -

- aborted RNA synthesis and returned to open complexes.

### EXAMPLE 5

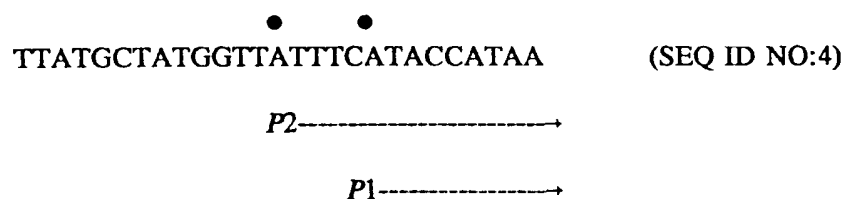
#### Differences in the behavior of the *gal* promoters

5                   Although it is difficult to precisely determine the microscopic constants,  $K_s$  and  $k_r$ , for *P1* and *P2* because of overlapping and the competitive nature of the two promoters, *P1* appears to be defective in  $K_s$  and *P2* is limiting in  $k_r$  (29). But, open complexes were formed at the two promoters at comparable efficiencies in the absence of cAMP. Herbert et al. (29) reported that the ratio of  $K_s \times k_r$  for *P1* and *P2* is roughly 1:1, whereas the same ratio determined by Goodrich and McClure (30) is 1:3. In the presence of cAMP, the ratio changed to greater than 50:1 in favor of *P1*. In the direct transcription assay of the invention, it was found that in the absence of cAMP, the total amount of aborted transcripts made by *P2* was at least 10-fold higher than those made by *P1*. This measurement was obtained by determining the molar ratios of each transcript after correcting for the number of uridine residues present in each species of RNA observed in Figure 4A and 4B. Synthesis of aborted transcripts from *P1* and *P2* is unequal in favor of *P2* in the absence of cAMP, but formation of the corresponding open complex is not unequal, suggesting that the *P1* promoter is inefficient in the formation of the first phosphodiester bond. By comparing the molar ratios of the aborted to the full-length transcripts, it was found in the present invention that greater than 90% of transcription that initiated at *P1* cleared the promoter, resulting in the synthesis of full-length transcripts in the presence of 100  $\mu$ M ATP, GTP, CTP, and 10  $\mu$ M UTP. Nierman and Chamberlin (31) have shown that productive elongation of RNA chain is a much more frequent event than dinucleotide formation in the A1 promoter phage T7, and Levin et al. (6) have subsequently demonstrated that the synthesis of 6-, 7-, and 8-mer RNA occurs at about the same rate as the production of full-length transcripts from the same promoter. However, greater than 90% of the transcription that initiated at *P2* aborted. Thus, RNA polymerase idled a lot more at *P2* than at *P1*. This demonstrated that the enzyme encountered a barrier to promoter clearance at the *P2* promoter. The presence of cAMP, which inhibited

- 24 -

- both aborted and full length RNA synthesis from *P2*, and enhanced RNA synthesis from *P1*, did not alter the ratio of aborted to full length transcripts from *P1*. Thus, cAMP acted at a step prior to the first phosphodiester bond formation.

The aborted initiation assays mentioned above have shown that cAMP stimulated *P1* by enhancing open complex formation (29), presumably by increasing  $K_a$  (30). The cAMP dependent *E. coli lacP1* promoter as well as its cAMP independent *UV5* variants (5) and the lambda  $P_r$  promoter (3) can initiate the synthesis of significant amounts of aborted transcripts along with elongated transcripts at high concentrations of nucleoside triphosphates. It has been suggested that the nucleotide sequence of the downstream region of a promoter contributes to the lifetime of an idling initial transcribing complex (6, 54) and that the rates of open complex formation need not correlate with rates of synthesis of full-length transcripts (55). The nucleotide sequences around the start sites of the *P2* and *P1* transcripts are:



Although it is possible that the trimer AUU made from *P2* may arise under the limiting UTP concentration used in the present invention, the hexamer from *P2* does not terminate at a uridine-rich area and thus must abort because of a kinetic or energetic barrier. It is also conceivable that the presence of the *P1* promoter 5 bp downstream from *P2* makes the RNA polymerase stall while traversing through *P1*.

The effect of *P1* on transcription from *P2* was tested by using a *P1* mutant *gal* DNA minicircle template made from plasmid pSA515. The mutant carried an A:T to T:A transversion at the -11 position of *P1*, which primarily affected *P1* activity (16). The results of synthesis of full-length RNA from this mutant as analyzed on an 8% gel are shown in Figure 6A. In the absence of cAMP, the mutation reduced *P2* activity less than two-fold, but inhibited *P1*

- 25 -

activity severely; the full-length *P1* transcript in the *P1* mutant was less than 10% of wild type. However, the response to cAMP•CRP was unaltered. In the presence of 20  $\mu$ M cAMP, *P2* RNA was undetectable and the residual *P1* RNA reached maximum as was shown above for wild type. The synthesis of aborted transcripts was shown in 25% gels (Figure 6B). At cAMP concentrations of 0 and 0.2  $\mu$ M, the full-length *P1* transcript was undetectable in the *P1* mutant; under these conditions, the pattern and amount of aborted transcripts from *P2* were identical to those from wild type *gal* DNA, suggesting that the generation of aborted transcripts from *P2* was independent of at least *P1* activity. Although it remains to be seen whether or not the actual DNA sequence of the *P1* region enhances idling at *P2*, as suggested by Levin et al. (6), the idling of RNA polymerase and aborted cycling of transcription from *P2* were specific and were likely to be an intrinsic property of the *P2* promoter. The biological significance of a barrier to promoter clearance, if any, in the initiation of transcription at *galP2* remains to be demonstrated. It has been suggested that the *malT* promoter of *E. coli* may be regulated at this level by cAMP (32).

In summary, the present invention showed that the *P1* promoter is limiting in closed complex and initial transcribing complex formation, while the *P2* promoter is slow in isomerization and promoter clearance. In spite of the intrinsic and major differences in the nature of limiting steps of RNA polymerase for transcription initiation at the two *gal* promoters, complete transcripts were synthesized at about equal efficiencies from *P1* and *P2* in the absence of cAMP. The transcripts were synthesized almost exclusively from *P1*, at 3.5-fold stimulation, in saturating concentrations of cAMP (Figure 4C) in the transcription assays of the present invention. Since the supercoiled minicircle transcription system of the present invention duplicated the *in vivo* situation (19), it provides a simple and physiologically appropriate system for studying all of the steps, including idling and promoter clearance, for any promoter and its regulatory aspects.

35

- 26 -

**EXAMPLE 6**Repression of *gal* transcription by wild type and mutant Lac repressors

Repression of *gal* transcription was studied using minicircle *gal* DNA template (pSA510), containing *lac* operator sequences in place of the *gal* operators at  $O_E$  and  $O_I$  ( $O_E^+$  and  $O_I^+$ , respectively) (37). Two types of Lac repressor were used to study repression: wild type  $LacI^+$  and mutant  $Lac^{adi}$  repressors. Wild type Lac repressor forms a tetramer and represses *gal* operon containing the *lac* operators *in vivo*, whereas the mutant repressor is a non-tetrameric protein and fails to repress under similar conditions (38). The mutant repressor exhibits normal binding to operators but fails to carry out the association of the dimer units into a tetramer as judged by electron microscopy (38, 44) and electrophoresis of operator-repressor complexes (45). Figure 7 shows the effect of  $LacI^+$  repressor on *gal* transcription *in vitro* in the absence (lanes 1-8) and presence (lanes 9-16) of 0.2 mM cAMP. Increasing concentrations of  $LacI^+$  protein repressed *P1* in the absence and presence of cAMP, and *P2* in the absence of cAMP (Figure 7, lanes 1-4 and 9-12). CRP was present at 100 nM. Quantification by direct scanning of the radioactivity of the full length transcripts in the gels is shown graphically in Figure 7C.  $LacI^+$  repressor was able to repress *gal* transcription greater than 95%. Similar repression of the aborted transcripts made from the *gal* promoters by the presence of  $LacI^+$  repressor was also observed (Figure 7B). However, in the absence of cAMP,  $LacI^{adi}$  repressed *P1* only about 75% and *P2* less than 10% (Figures 8A and 8B; lanes 1-4). In the presence of cAMP, *P1* was repressed by  $LacI^{adi}$  by 75% (lanes 9-12). The results are shown quantitatively in Fig. 8C. The inhibition of *P2* activity by cAMP•CRP was relieved slightly by  $LacI^{adi}$ . This was more apparent for the aborted transcripts (trimer and hexamer) from *P2* (Figure 8B, lanes 1-4 and 9-12). The repression of *P1* and *P2* by  $LacI^+$  and the weak repression of *P1* by  $LacI^{adi}$  were totally relieved by the addition of 1 mM IPTG, which inactivates Lac repressors, showing that the inhibition of transcription was Lac repressor specific (lanes 5-8 and 13-16 in Figures 7 and 8). The repression by  $LacI^+$  was almost complete in every case with half maximal repression obtained at approximately 10 nM repressor (Figure 7C). The maximal repression of *P1*

- 27 -

- transcript by Lac<sup>adi</sup> of only 75 % required very high repressor concentration (90 mM) with a half-maximal concentration of about 60 nM (Figure 8C). Further increase of LacI<sup>adi</sup> repressor concentration had a nonspecific inhibitory effect: the inducer (IPTG) failed to derepress.

5

#### Property of Gal repressor

The effect of Gal repressor (GalR<sup>+</sup>) on transcription was investigated *in vitro* in the absence or presence of 0.2 mM cAMP using the minicircle DNA template containing wild type *gal* operators (pSA509, see Figure 3) at O<sub>E</sub> and O<sub>I</sub> (O<sub>E</sub><sup>G</sup> and O<sub>I</sub><sup>G</sup>, respectively). CRP was present at 100 nM. The synthesis of both full length and aborted *gal* RNA was investigated under these conditions. Quantification of the full length transcripts is shown in Figure 9A. In the presence of cAMP, when only P1 is transcribed, approximately 80 nM GalR<sup>+</sup> was able to reduce the P1 transcripts only about 80% (right hand panel). P2 transcripts, both aborted and full length, were undetectable in the presence of cAMP but increased slightly with increasing concentration of GalR<sup>+</sup>. In the absence of cAMP, the RNA band from P1 was reduced by GalR<sup>+</sup> about 60% (left hand panel). Thus, a significant amount of *gal* transcription from P1 was maintained in the presence of saturating concentrations of GalR<sup>+</sup> irrespective of the presence of cAMP. P2 activity was not repressed by GalR<sup>+</sup>. In fact, P2 activity, both in the absence and in the presence of cAMP, increased slightly. The poor repressive effects of GalR<sup>+</sup> on P1 in the absence and presence of cAMP were abolished by the addition of 10 mM D-galactose.

25

### EXAMPLE 7

#### Mutant operators

30

The amount of partial repression by GalR<sup>+</sup> repressor described above was roughly equivalent to that obtained *in vivo* with an intact O<sub>E</sub> (O<sub>E</sub><sup>G</sup>) and mutant O<sub>I</sub> (O<sub>I</sub><sup>G</sup>) (37). In order to investigate whether the partial repression in the purified system was due to an interaction between GalR<sup>+</sup> and O<sub>E</sub> (O<sub>E</sub><sup>G</sup>) alone, DNA minicircle templates with mutations at either O<sub>I</sub> or O<sub>E</sub> were used to determine the

35

- 28 -

nature of repression with GalR<sup>+</sup> *in vitro*. Figure 9B shows the effect of GalR<sup>+</sup> concentrations on the repression of transcription using minicircle DNA (pSA511) with the genotype O<sub>E</sub><sup>c</sup> O<sub>I</sub><sup>c</sup>. This mutant DNA (pSA511) and the wild type DNA (pSA509) provided results that were more or less indistinguishable from each other: P1 transcripts were repressed about 80% at about 80 nM GalR<sup>+</sup> both in the presence or absence of cAMP, and P2 was not repressed in the absence of cAMP. The *gal* O<sub>E</sub><sup>c</sup> O<sub>I</sub><sup>c</sup> DNA (pSA512), showed no discernable effect on P1 and P2 (Figure 9C). Even in the presence of very high GalR<sup>+</sup> concentrations, very little change of *gal* RNA synthesis was observed either in the absence or in the presence of cAMP.

## LIST OF REFERENCES

1. Walker, G., Zillig, W., Palm, P. & Fuchs, E. (1967) *Eur. J. Biochem.* **3**, 194-201.
2. Chamberlin, M. J. (1974) *Ann. Rev. Biochem.* **43**, 771-775.
3. McClure, W. R., Cech, C. L. & Johnston, D. E. (1978) *J. Biol. Chem.* **253**, 8941-8948.
4. Carpousis, A. J. & Gralla, J. D. (1980) *Biochemistry* **19**, 3245-3253.
5. Munson, L. M. & Reznikoff, W. S. (1981) *Biochemistry* **20**, 2081-2085.
6. Levin, J. R., Krummel, B. & Chamberlin, M. J. (1987) *J. Mol. Biol.* **196**, 85-100.
7. Krummel, B. & Chamberlin, M. J. (1989) *Biochemistry* **28**, 7827-7842.
8. Adhya, S. (1989) *Ann Rev. Genet.* **23**, 227-250.
9. McClure, W. R. (1985) *Ann. Rev. Biochem.* **54**, 171-204.
10. Kramer, H., Amouyal, M., Nordheim, A., Muller-Hill, B. (1988) *EMBO J.* **7**, 547-556.
11. Meiklejohn, A. L. & Gralla, J. D. (1989) *J. Mol. Biol.* **207**, 661-673.
12. Hahn, S., Hendrickson, W. & Schleif, R. (1986) *J. Mol. Biol.* **188**, 355-367.
13. Campbell, A. M. (1962) *Adv. Genet.* **11**, 101-145.

14. Nash, H. (1981) *Ann. Rev. Genet.* **15**, 143-167.
15. Bliska, J. B. & Cozzarelli, N. R. (1987) *J. Mol. Biol.* **194**, 205-218.
16. Musso, R. E., DiLauro, R., Adhya, S. & de Crombrughe, B. (1977) *Cell* **12**, 847-854.
- 5 17. Adhya, S. & Miller, W. (1979) *Nature* **279**, 492-494.
18. Aiba, H., Adhya, S. & de Crombrughe, B. (1981) *J. Biol. Chem.* **265**, 11905-11910.
19. Irani, M., Musso, R. & Adhya, S. (1987) *J. Bacteriol.* **171**, 1623-1630.
- 10 20. Mizuuchi, M. & Mizuuchi, K. (1985) *Nucleic Acid Res.* **13**, 1193-1208.
21. Squires, C., Krainer, A., Barry, G., Shen, W.-F. & Squires, C.L. (1981) *Nucleic Acid Res.* **9**, 16827-16840.
- 15 22. Hsu, D.L., Ross, W. Landy, A. (1980) *Nature* **285**, 85-91.
23. Irani, M., Orosz, L. & Adhya, S. (1983) *Cell* **32**, 783-788.
24. Guarneros, G. & Echols, H. (1973) *Virology* **52**, 30-38.
- 20 25. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. eds. (1989) *Current Protocols in Molecular Biology*. (Green/Wiley-Interscience, New York).
26. Echols, H. (1970) *J. Mol. Biol.* **47**, 575-583.
27. Kravchenko, V. V., Vassilenko, S. K. & Grachev, M. A. (1979) *Gene* **7**, 181-195.
- 25 28. Kur, J., Hasan, N. & Szybalski, W. (1989) *Virology* **168**, 236-244.
29. Herbert, M., Kolb, A. & Buc, H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2807-2811.
- 30 30. Goodrich, J. A. & McClure, W. R. (1992) *J. Mol. Biol.* **224**, 15-29.
31. Nierman, W. C. & Chamberlin, M. (1979) *J. Biol. Chem.* **254**, 7921-7926.
- 35 32. Menendez, M., Kolb, A. & Buc, H. (1987) *EMBO J.* **6**, 4227-4234.

- 30 -

33. Nash, H. (1974) *Virology* 57, 207.
34. Matthews, K. S. (1992) *Microbiol. Rev.* 56, 123-136,
35. Fritz, H.-J., Bicknase, H., Gleumes, B., Heibach, C., Roshl, S., & Ehring, R. (1983) *EMBO J.* 2, 2129-2135.
- 5 36. Majumdar, A. & Adhya, S. (1984) *Proc. Natl. Acad. Sci. USA* 81, 6100-6104.
37. Haber, R. & Adhya, S. (1988) *Proc. Natl. Acad. Sci. USA* 85, 9683-9687.
- 10 38. Mandal, N., Su, W., Haber, R., Adhya, S. & Echols, H. (1990) *Genes Dev.* 4, 410-418.
39. Adhya, S. & Majumdar, A. (1987) in *RNA Polymerase and Regulation of Transcription* (Reznikoff, W., Burgess, R., Dahlberg, J., Gross, C., Record, T. & Wickens, M. eds.): pp. 129-135.
- 15 40. Kuhnke, G., Theres, C., Fritz, H.-J. & Ehring, R. (1989) *EMBO J.* 8, 1247-1255.
41. Browiec, J. A., Zhang, L., Sasse-Dwight, S. & Gralla, J. D. (1987) *J. Mol. Biol.* 196, 101-111.
- 20 42. Lobell, R. B. & Schleif, R. F. (1991) *Science* 250, 528-532.
43. Majumdar, A., Rudikoff, S. & Adhya, S. (1987) *J. Biol. Chem.* 262, 2326-2331.
44. Lehming, N., Sartorius, J., Oehler, S., von Wilcken-Bergmann, B. & Muller-Hill, B. (1988). *Proc. Natl. Acad. Sci. USA* 85, 7947-7951.
- 25 45. Brenowitz, M., Mandal, N. Pickar, A., Jamison, E. & Adhya, S. (1991) *J. Biol. Chem.* 266, 1281-1288.
46. Kramer, H., Niemoller, M., Amouyal, M., Revet, B., Wilcken-Bergman, B. & Muller-Hill, B. (1987) *EMBO J.* 6, 1481-1491.
- 30 47. Nick, H. & Gilbert, W. (1985) *Nature* 313, 795-798.
48. de Crombrughe, B., Chen, B., Anderson, W., Nissley, P., Gottesman, M., Pastan, I. & Perlman, R. (1971) *Nature-New Biology* 231, 139-142.
- 35 49. Straney, S. B. & Crothers, D. M. (1987) *Cell* 51, 699-707.



- 31 -

50. Lee, J. & Goldfard, A. (1991) *Cell* **66**, 793-798.
51. Brenowitz, M., Jamison, E., Majumdar, A. & Adhya, S. (1990) *Biochemistry* **29**, 3374-3383.
52. Kuhnke, G., Kraus, A., Heibach, C., Gieske, U., Fitz, H.-J. & Ehring, R. (1986) *EMBO J.* **5**, 167-173.
53. Alberti, S., Oehler, S., Wilcken-Bergmann, B., Kramer, H. & Muller-Hill, B. (1991) *New Biologists* **3**, 57-62.
54. Kammerer, W., Deuschle, U., Gentz, R., Bujard, H. (1986) *EMBO J.* **5**, 2995-3000.
55. Knaus, R. and Bujard, H. (1990) In: *Nucleic Acids and Molecular Biology*, Eds: F. Eckstein and D.M.J. Lilly, **4**, pp. 110-122.

- 32 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANTS: ADHYA, SANKAR L.; CHOY, HYON E.

5

(ii) TITLE OF INVENTION: SUPERCOILED MINICIRCLE DNA  
AS A UNITARY PROMOTER VECTOR

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

10

(A) ADDRESSEE: MORGAN & FINNEGAN  
(B) STREET: 345 PARK AVENUE  
(C) CITY: NEW YORK  
(D) STATE: NEW YORK  
(E) COUNTRY: USA  
(F) ZIP: 10154

15

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: FLOPPY DISK  
(B) COMPUTER: IBM PC COMPATIBLE  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: WORDPERFECT 5.1

20

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE: 16-OCT-1992  
(C) CLASSIFICATION:

25

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

30

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: LESLIE SERUNIAN AND/OR MARY J. MORRY  
(B) REGISTRATION NUMBER: 35,353 AND/OR 34,398  
(C) REFERENCE/DOCKET NUMBER: 2026-4041

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212)758-4800  
(B) TELEFAX: (212)751-6849

35

(2) INFORMATION FOR SEQ ID NO:1:

- 33 -

- ° (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 54  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: UNKNOWN
- 5 (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: E. COLI  
(B) STRAIN:  
10 (C) INDIVIDUAL ISOLATE:  
(D) DEVELOPMENTAL STAGE:  
(E) HAPLOTYPE:  
(F) TISSUE TYPE:  
(G) CELL TYPE:  
(H) CELL LINE:  
15 (I) ORGANELLE:
- (ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
20 (D) OTHER INFORMATION: THE 54 BASE PAIR RHO  
INDEPENDENT TRANSCRIPTION TERMINATOR OF  
THE RPO-BC GENE OF E. COLI
- (xi) SEQUENCE DESCRIPTIONS:SEQ ID NO:1:
- GCAAATAACG TAAAAACCCG CTCGGCGGG TTTTTTATG 40
- 25 GGGGGAGTTT AGGG 54
- (3) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 156  
30 (B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: UNKNOWN
- (ii) MOLECULE TYPE: DNA
- 35 (iii) HYPOTHETICAL: YES

- 34 -

- ° (vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE:  
(D) DEVELOPMENTAL STAGE:  
(E) HAPLOTYPE:  
5 (F) TISSUE TYPE:  
(G) CELL TYPE:  
(H) CELL LINE:  
(I) ORGANELLE:
- 10 (ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION: THE 156 BASE PAIR  
SEQUENCE FROM THE BACTERIAL ATT SITE TO  
THE TRANSCRIPTION TERMINATOR IN PLASMID  
PSA508
- 15 (xi) SEQUENCE DESCRIPTIONS:SEQ ID NO:2:
- |  |     |
|--|-----|
| AATTTCTAGA CCTTCCCGTT TCGCTCAAGT TAGTATAAAA    | 40  |
| AAGCAGGCTT CAACCGAGCT CGTCGACCCG GGTACCGAGC    | 80  |
| 20 ATGCATGAAT TCGTCTGCAG CAAATAACGT AAAAACCCGC | 120 |
| TTCGGCGGGT TTTTTTATGG GGGGAGTTTA GGGATC        | 156 |
- (4) INFORMATION FOR SEQ ID NO:3:
- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16  
(B) TYPE: NUCLEIC ACID  
(C) STRANDEDNESS: SINGLE  
(D) TOPOLOGY: UNKNOWN
- 30 (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
35 (C) INDIVIDUAL ISOLATE:

- 35 -

- °
- (D) DEVELOPMENTAL STAGE:
  - (E) HAPLOTYPE:
  - (F) TISSUE TYPE:
  - (G) CELL TYPE:
  - (H) CELL LINE:
  - (I) ORGANELLE:
- 5
- (ix) FEATURE:
    - (A) NAME/KEY:
    - (B) LOCATION:
    - (C) IDENTIFICATION METHOD:
    - (D) OTHER INFORMATION: CONSENSUS LAC OPERATOR SEQUENCE
- 10
- (xi) SEQUENCE DESCRIPTIONS:SEQ ID NO:3:
- TTGTGAGCGC TCACAA
- 16
- (5) INFORMATION FOR SEQ ID NO:4:
- 15
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 27
    - (B) TYPE: NUCLEIC ACID
    - (C) STRANDEDNESS: SINGLE
    - (D) TOPOLOGY: UNKNOWN
- 20
- (ii) MOLECULE TYPE: DNA
  - (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE:
- 25
- (A) ORGANISM: E. COLI
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE:
  - (D) DEVELOPMENTAL STAGE:
  - (E) HAPLOTYPE:
  - (F) TISSUE TYPE:
  - (G) CELL TYPE:
  - (H) CELL LINE:
  - (I) ORGANELLE:
- 30
- (ix) FEATURE:
    - (A) NAME/KEY:
    - (B) LOCATION:
- 35

- 36 -

- (C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION: THE NUCLEOTIDE  
SEQUENCES AROUND THE START SITES OF THE  
P2 AND P1 TRANSCRIPTS OF THE GAL OPERON  
OF E. COLI

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTATGCTATG GTTATTTTCAT ACCATAA

27

10

15

20

25

30

35

- 37 -

## WHAT IS CLAIMED IS:

1. A DNA construct comprising, in the following order,

- a. bacterial attachment site *attB'OB*;
- b. a multiple cloning site;
- c. a transcription terminator; and
- d. lambda phage attachment site *attP'OP*;

which, when the construct is present in a host cell containing or modified to contain genes encoding the lambda Int protein and integration host factors (IHF), recombination occurs between the *attB'OB* and *attP'OP* sites upon expression of the lambda Int protein to generate a smaller supercoiled DNA minicircle comprising the multiple cloning site and the terminator.

2. The DNA construct according to claim 1, wherein the multiple cloning site further comprises a transcriptional promoter segment, wherein the transcription terminator is factor-independent, and wherein the smaller DNA minicircle comprises the terminator and the multiple cloning site containing the transcriptional promoter segment.

3. The DNA construct according to claim 2, wherein the transcriptional promoter segment comprises the *gal* promoter segment of the *gal* operon of *E. coli*.

4. The DNA construct according to claim 2, wherein the multiple cloning site further comprises one or more operator loci operably linked to the promoter segment.

5. The DNA construct according to claim 4, wherein the operator locus comprises the  $O_E$  operator of the *gal* operon of *E. coli*.

6. The DNA construct according to claim 4, wherein the operator locus comprises the  $O_I$  operator of the *gal* operon of *E. coli*.

- 38 -

7. The DNA construct according to claim 4, wherein the operator locus comprises the *lac* operator of the *lac* operon of *E. coli*.
8. The DNA construct according to claim 4, wherein the operator loci comprise the O<sub>E</sub> and the O<sub>I</sub> operators of the *gal* operon of *E. coli*.
9. A host cell containing the DNA construct according to claim 1.
10. The host cell according to claim 9, wherein the multiple cloning site further comprises a transcriptional promoter segment, wherein the terminator is factor-independent, and wherein the smaller minicircle DNA comprises the terminator and the multiple cloning site containing the transcriptional promoter.
11. The host cell according to claim 9 or claim 10, wherein the gene encoding the lambda Int protein is under the control of the temperature sensitive repressor *cI857*.
12. A method for analyzing transcripts from a promoter segment on a small DNA construct, comprising the steps of:
  - (a) introducing a DNA construct according to claim 2 into a host cell containing or modified to contain genes encoding the lambda Int protein and integration host factors (IHF);
  - (b) expressing the lambda Int protein, whereby a smaller DNA minicircle containing the multiple cloning site comprising the promoter segment and the transcription terminator is generated by recombination between the *attB'*OB and *attP'*OP sites in the DNA construct ;
  - (c) harvesting the cells;
  - (d) isolating the smaller DNA minicircle;
  - (e) preincubating the smaller DNA minicircle in a transcription



- 39 -

- mixture comprising a transcript-producing and compatible RNA polymerase;
- (f) adding ribonucleotide triphosphates to initiate a transcription reaction;
- 5 (g) terminating the transcription reaction; and
- (h) analyzing the RNA transcripts.

10 13. The method according to claim 12, further comprising adding host-cell compatible transcription initiation factors and/or transcription termination factors after the preincubation step (e), prior to or simultaneously with adding the ribonucleotide triphosphates of step (f).

15 14. The method according to claim 12, wherein the gene encoding the lambda Int protein is under the control of the temperature sensitive repressor *cI857* and wherein expression of the Int protein comprises the steps of:

- (a) growing the host cell in culture medium at 32°C to late log phase;
- (b) shifting the temperature of the culture medium to 42°C for 10 minutes; and
- 20 (c) shifting the temperature of the culture medium to 32°C for 30 minutes.

25 15. The host cell according to claim 9, said host cell being *E. coli*.

16. The method according to claim 12, said host cell of step (a) being *E. coli*.

30 17. Plasmid pSA508 according to Figure 1.

18. Plasmid pSA509 according to Figure 10.

35

1 / 18

**FIG. 1**

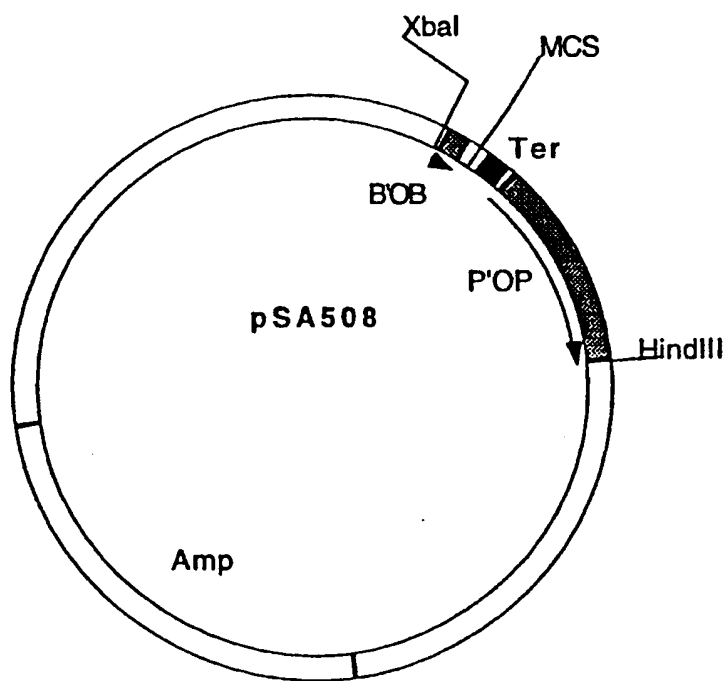
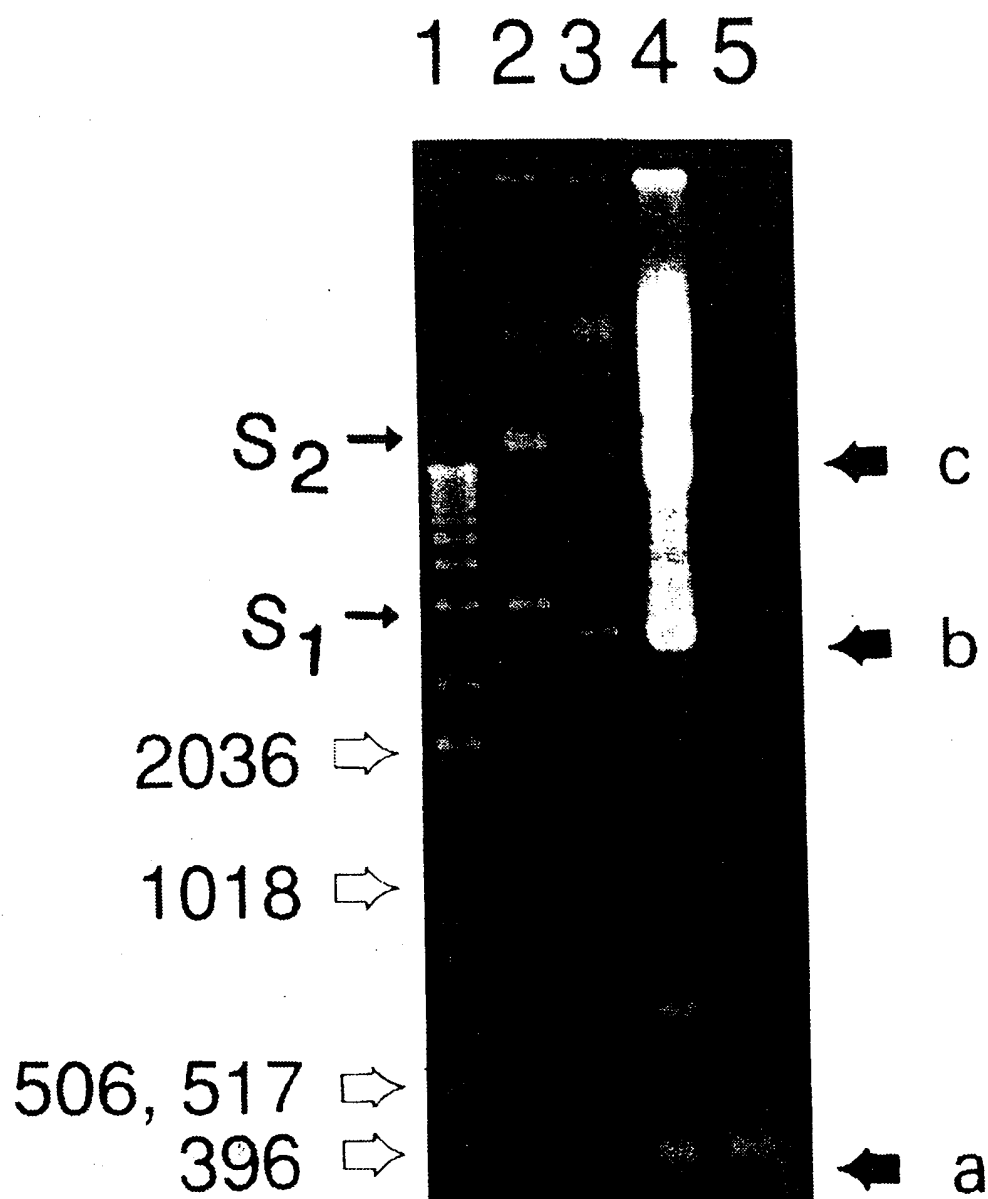
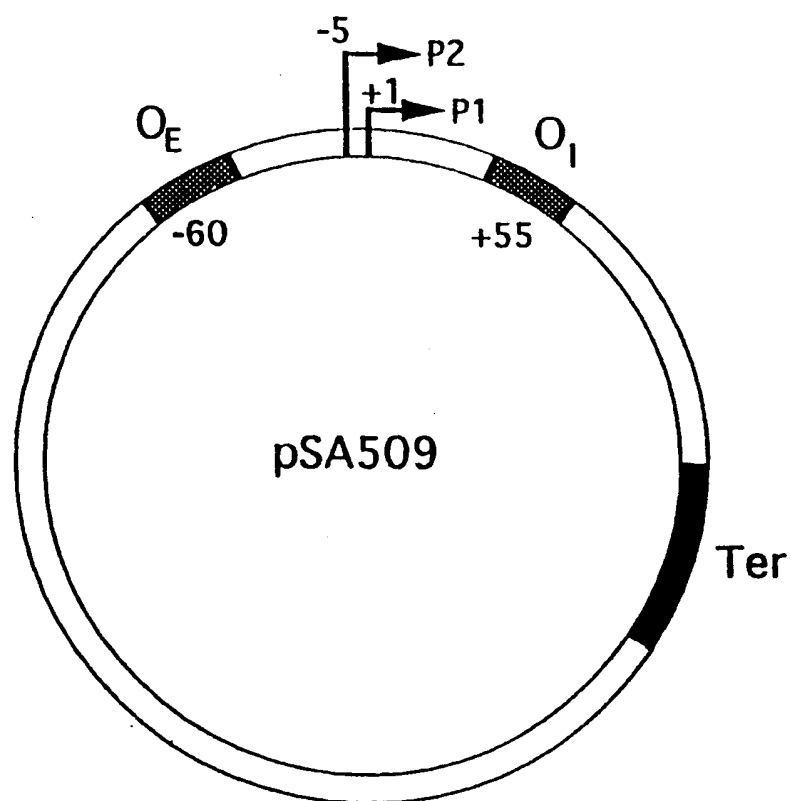


FIG . 2

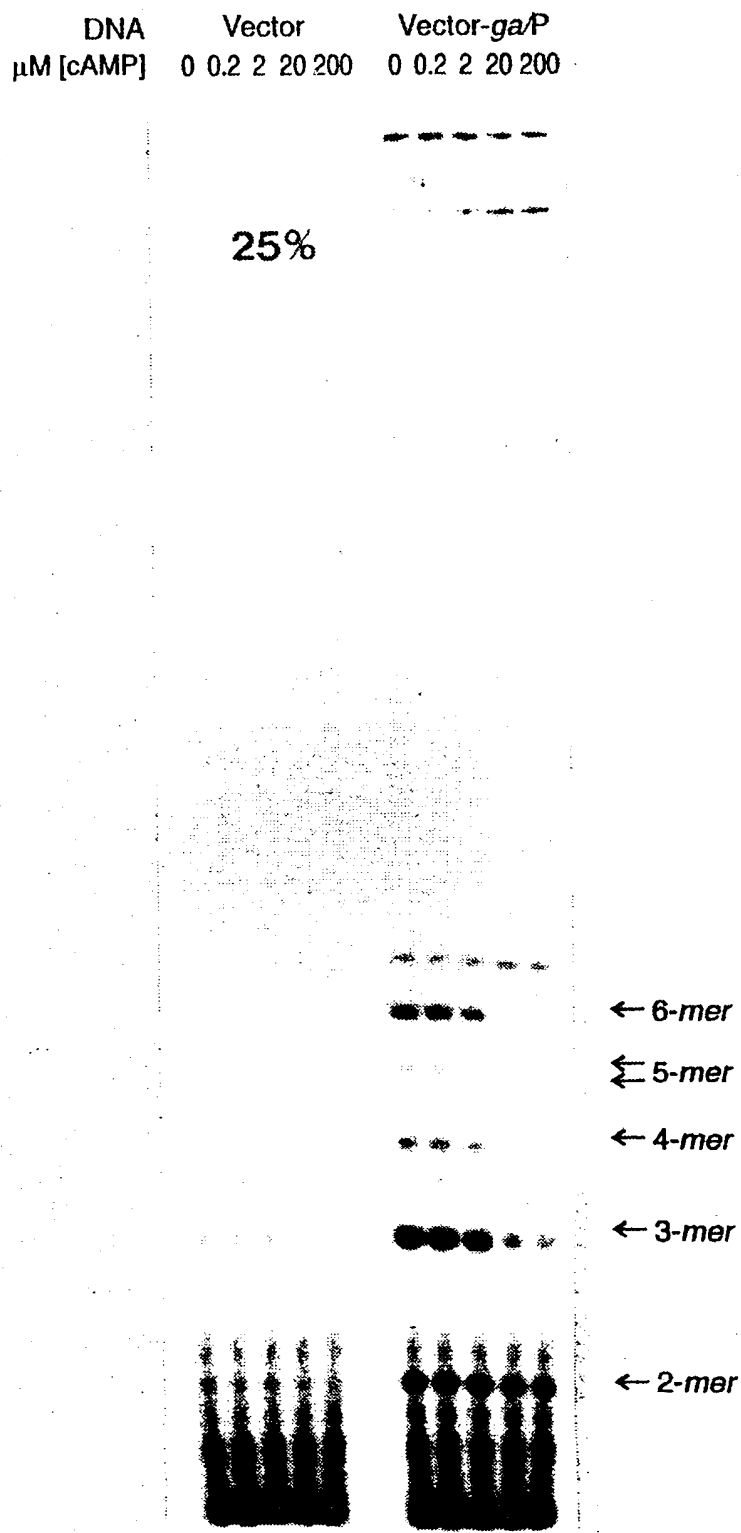


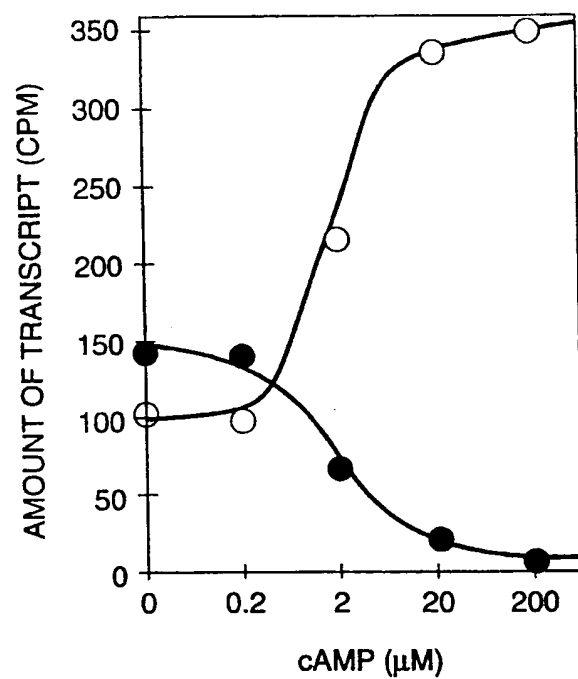
**FIG. 3**

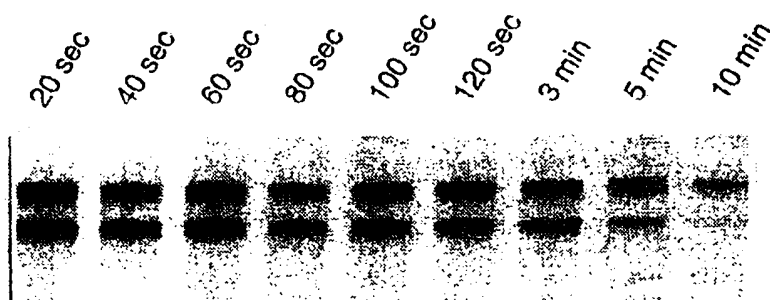
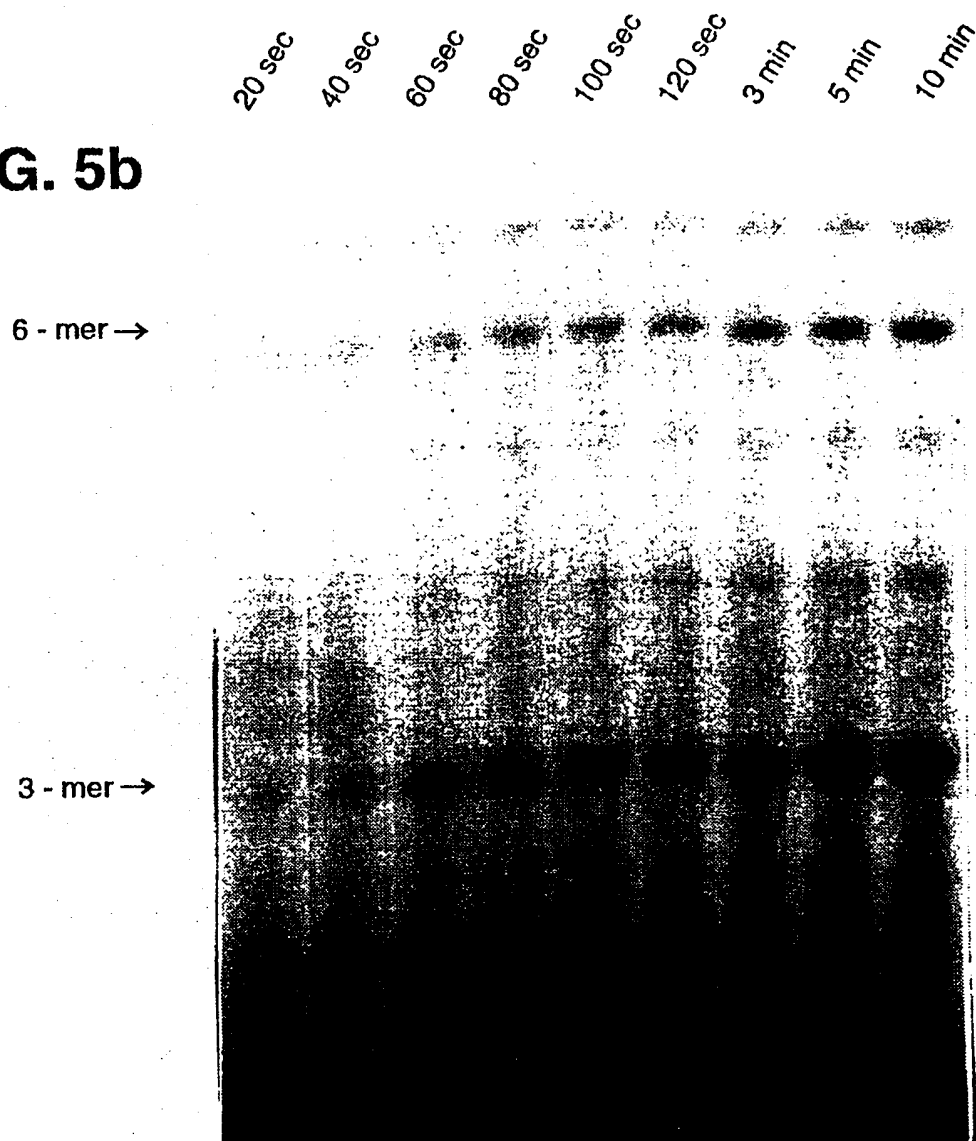
DNA $\mu$ M [cAMP]	Vector	Vector-ga/P
0	0.2	0
2	2	0.2
20	20	2
200	200	200

**FIG. 4a****8%** $\leftarrow$  6-mer $\leftarrow$  5-mer $\leftarrow$  4-mer $\leftarrow$  3-mer $\leftarrow$  2-mer

FIG. 4b



**FIG. 4c**

**FIG. 5a****FIG. 5b**



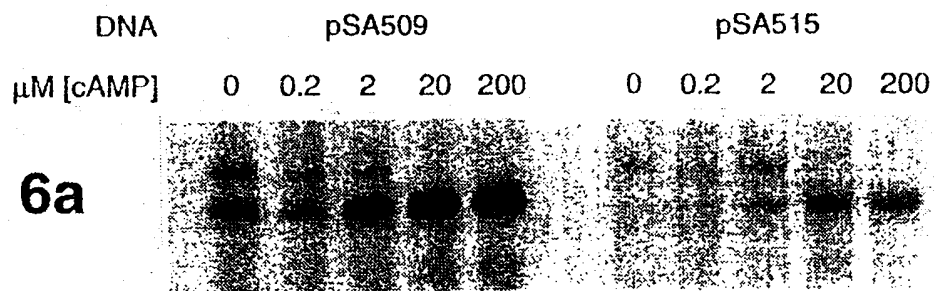
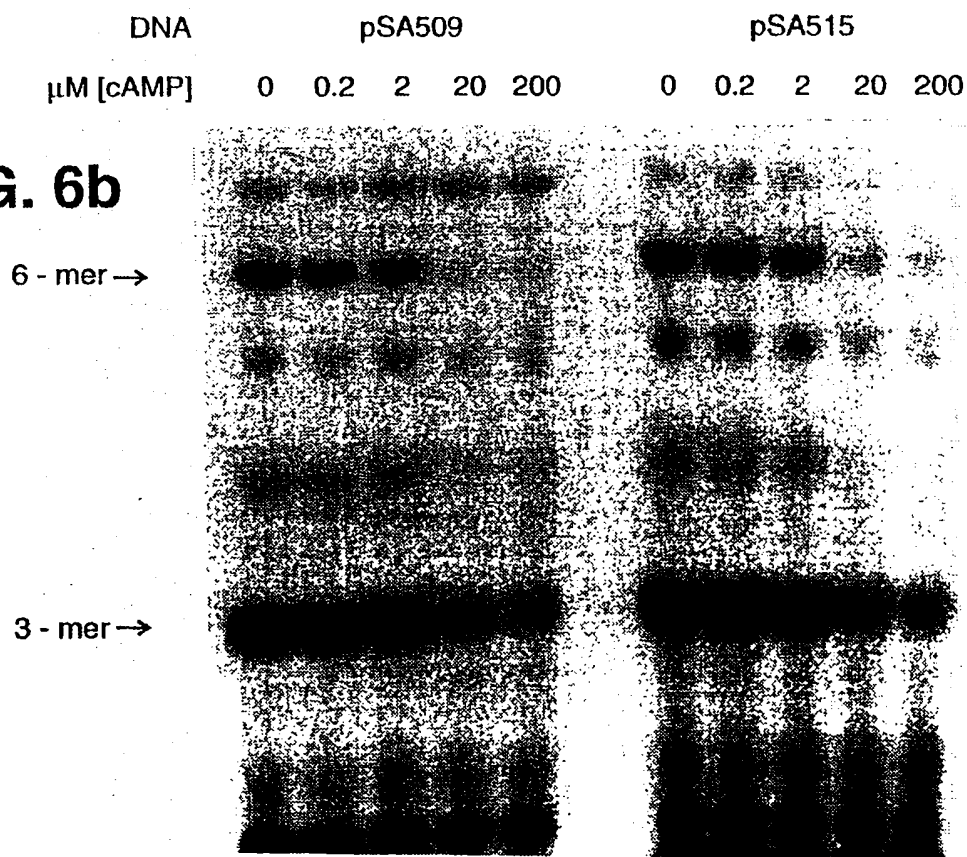
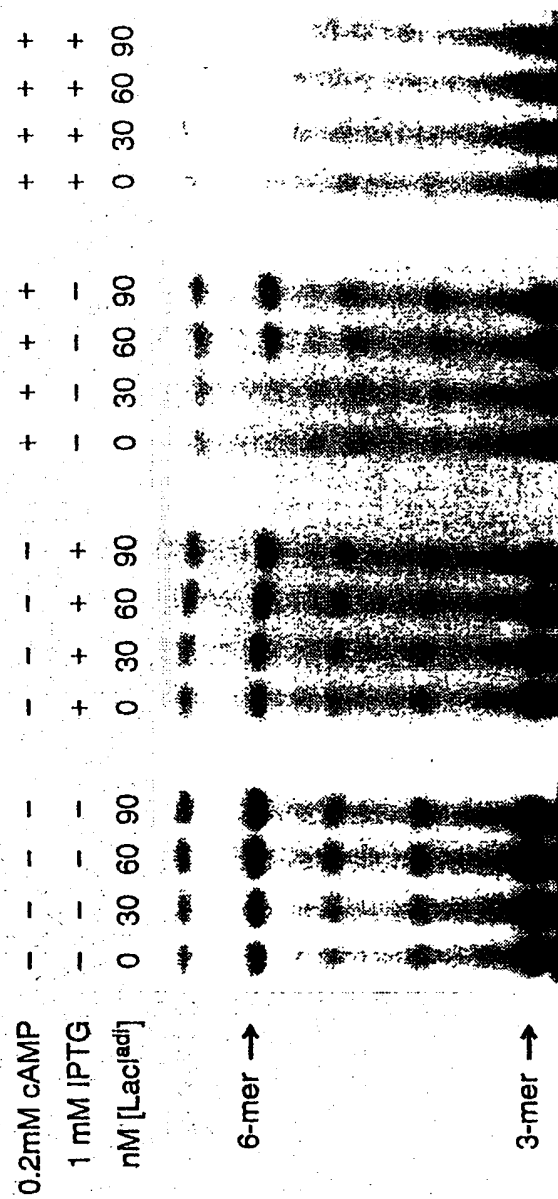
**FIG. 6a****FIG. 6b**



FIG. 7b



11 / 18

FIG. 7d

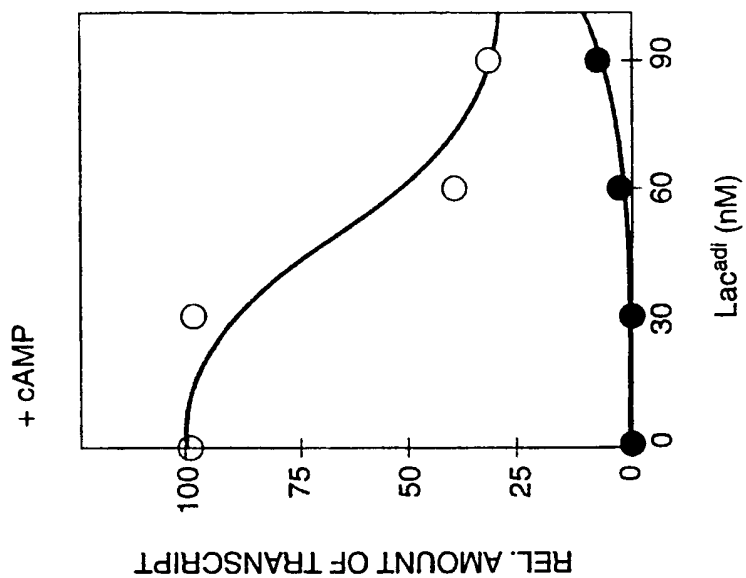


FIG. 7c

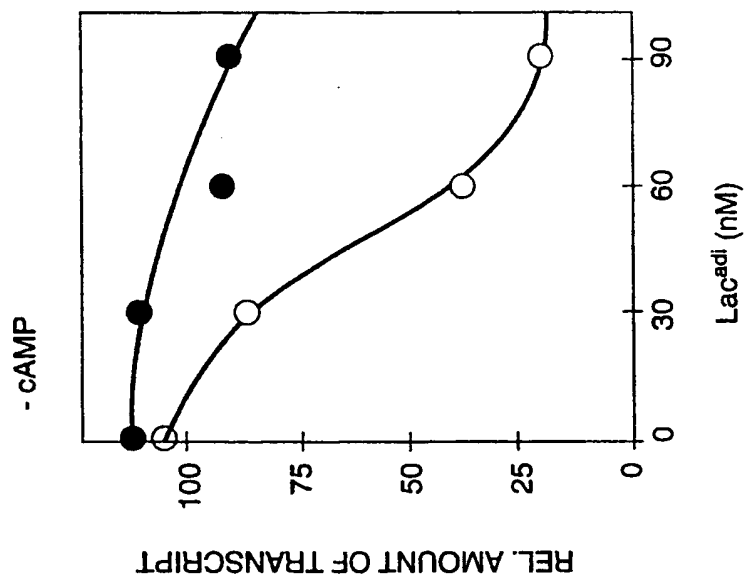






FIG. 8d

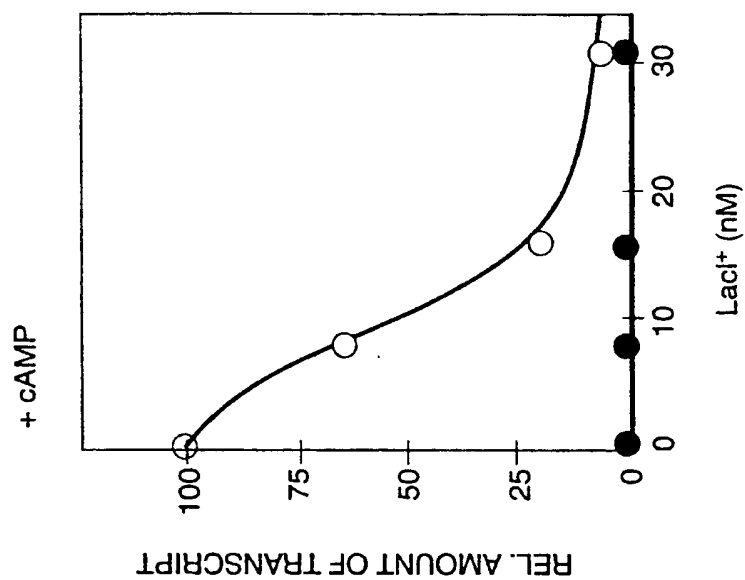


FIG. 8c

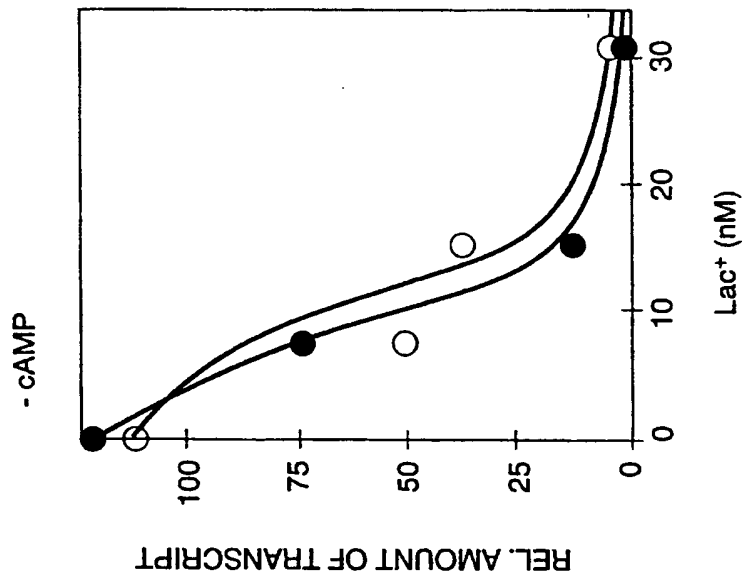


FIG. 9b

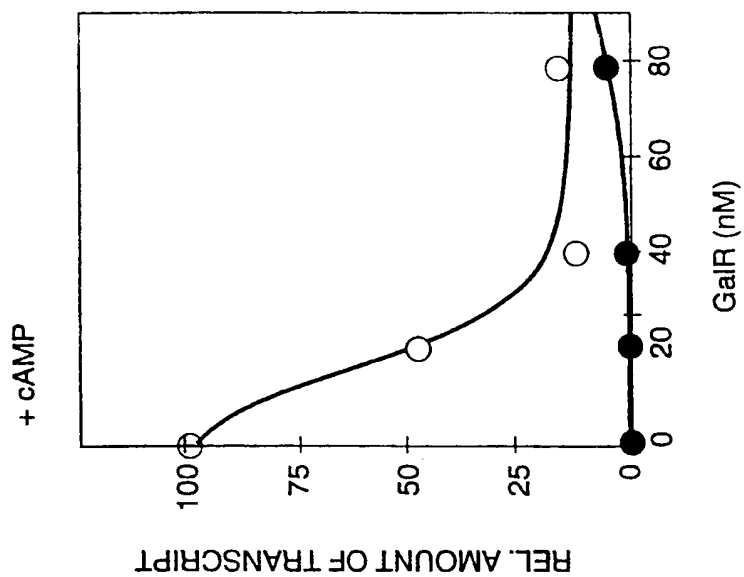
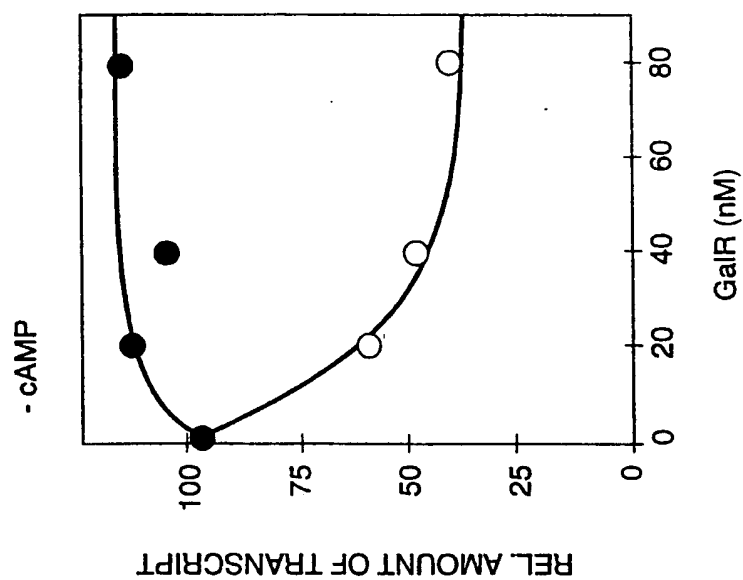


FIG. 9a





16 / 18

FIG. 9d

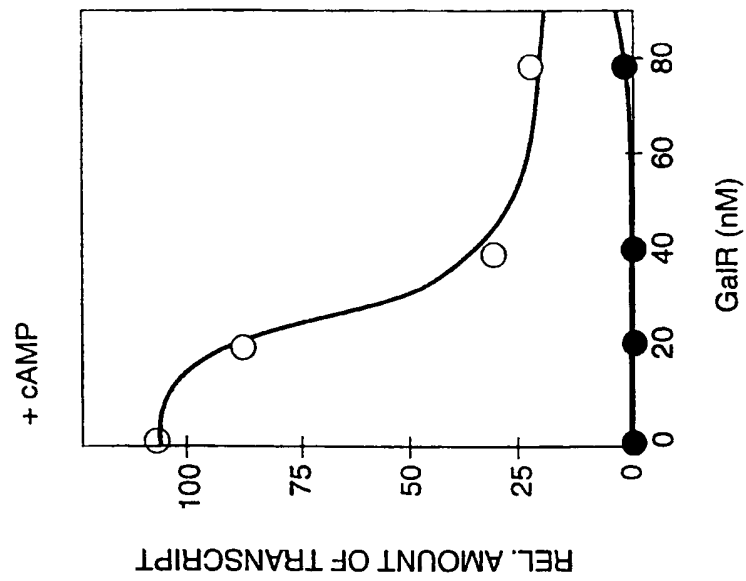
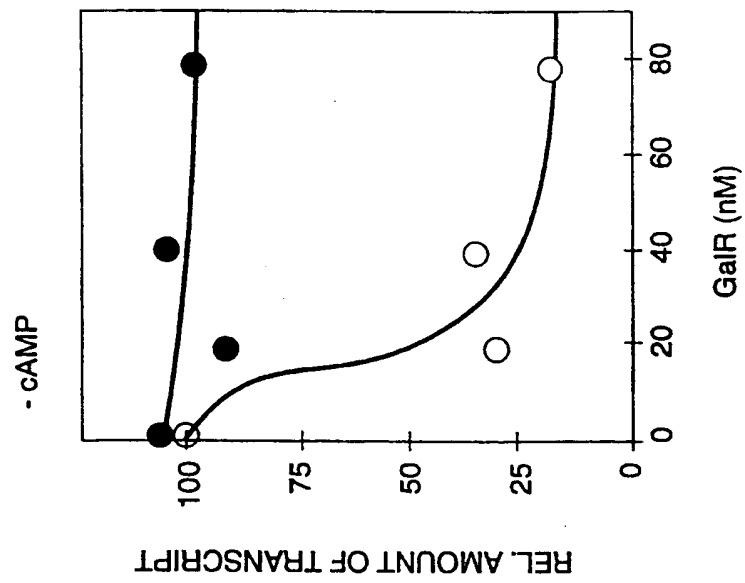


FIG. 9c



17 / 18

FIG. 9f

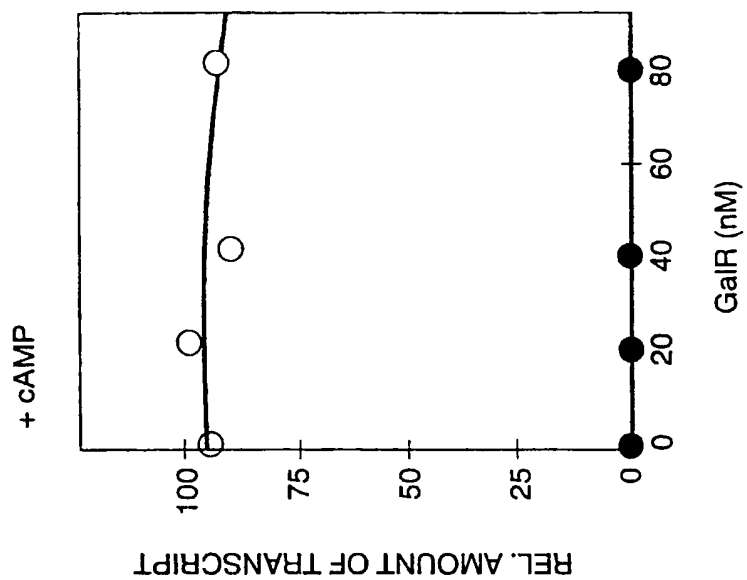
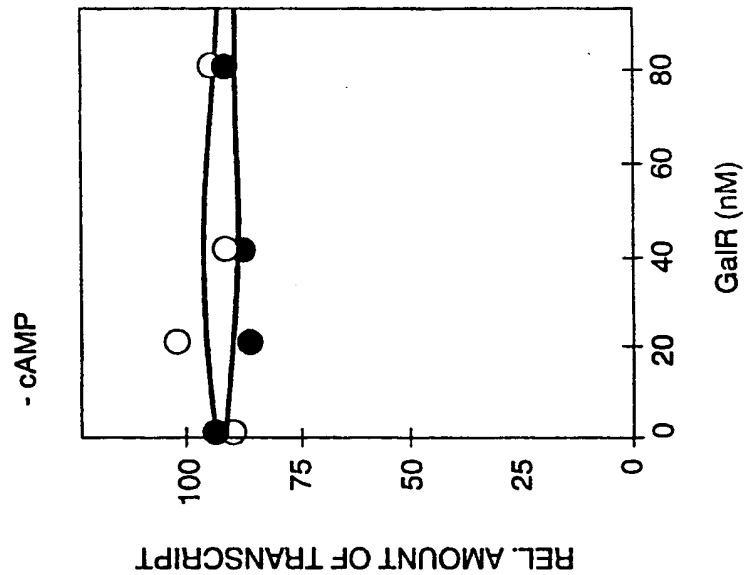
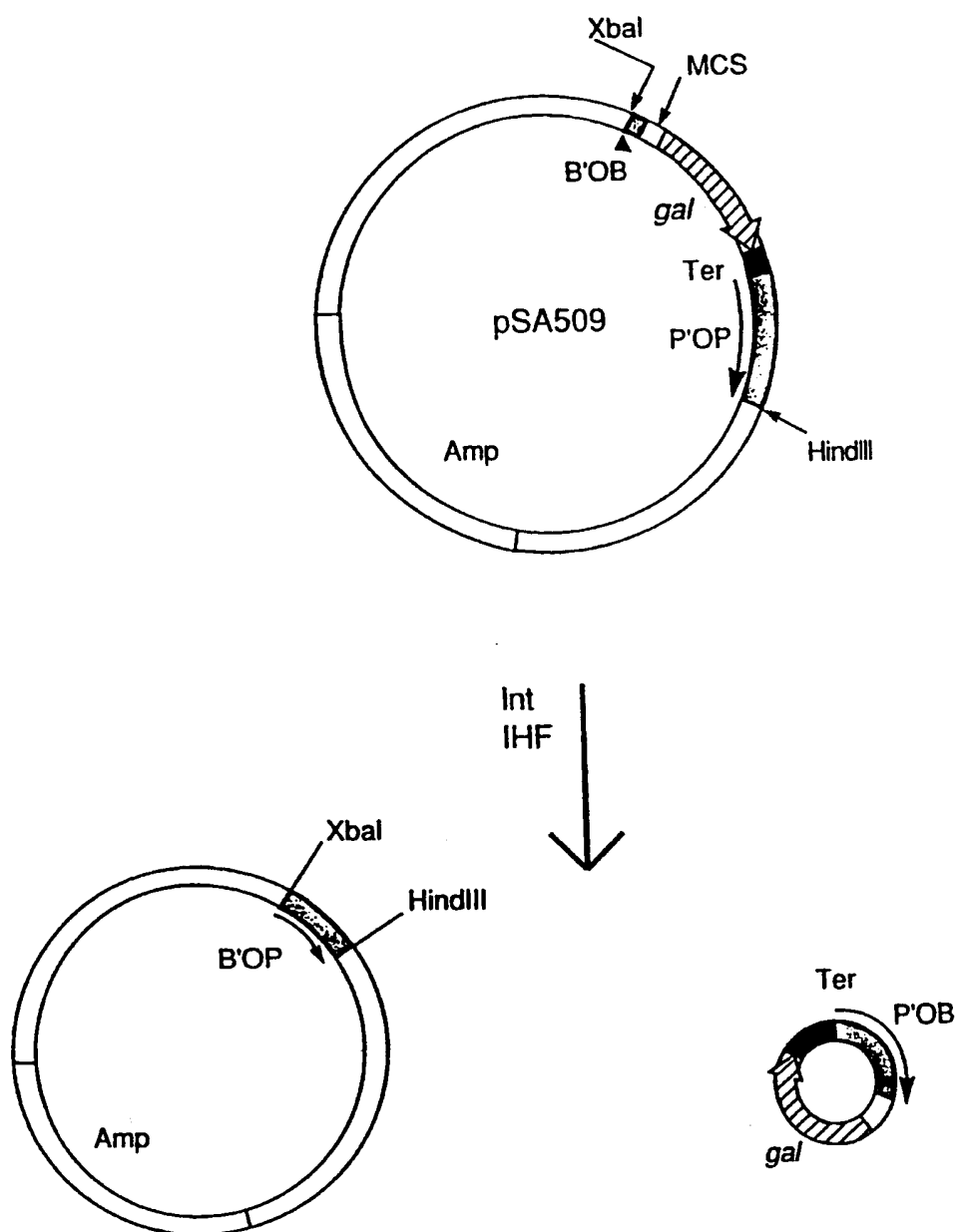


FIG. 9e



18 / 18

FIG. 10



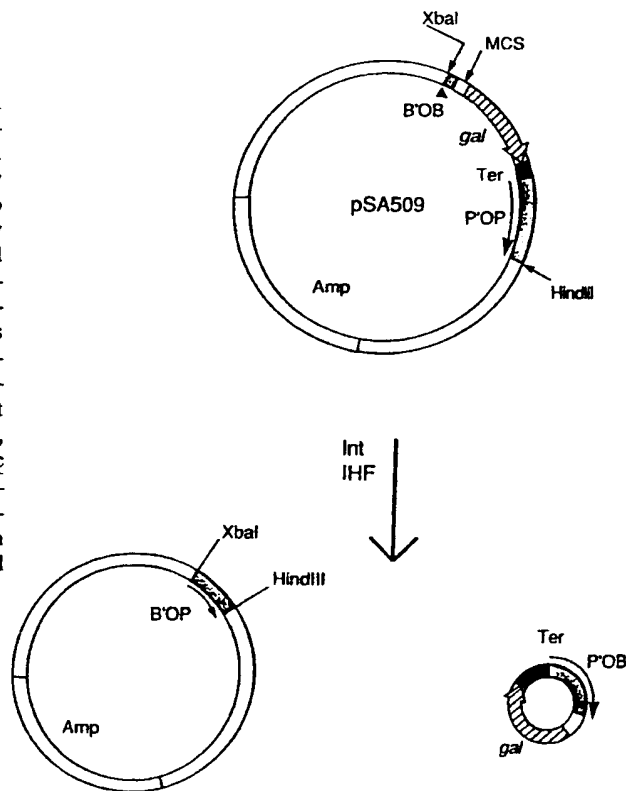


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> <b>C12N 15/10, 15/70, 1/21</b> <b>// (C12N 1/21, C12R 1:19)</b>	<b>A3</b>	<b>(11) International Publication Number:</b> <b>WO 94/09127</b> <b>(43) International Publication Date:</b> 28 April 1994 (28.04.94)
<b>(21) International Application Number:</b> PCT/US93/09918 <b>(22) International Filing Date:</b> 15 October 1993 (15.10.93) <b>(30) Priority data:</b> 07/963,328      16 October 1992 (16.10.92)      US <b>(71) Applicant:</b> THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Box OTT, Bethesda, MD 20892 (US). <b>(72) Inventors:</b> ADHYA, Sankar, L. ; 14400 Kings Grant Street, Gaithersburg, MD 20878 (US). CHOY, Hyon, E. ; 3 Pookshille, Bethesda, MD 20814 (US).	<b>(74) Agents:</b> FEILER, William, S. et al.; Morgan & Finnegan, 345 Park Avenue, New York, NY 10154 (US). <b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> <b>(88) Date of publication of the international search report:</b> 09 June 1994 (09.06.94)	

**(54) Title:** SUPERCOILED MINICIRCLE DNA AS A UNITARY PROMOTER VECTOR**(57) Abstract**

A system to construct supercoiled "minicircle" DNA containing a single promoter by the use of the integrative recombination system of bacteriophage lambda is described. The supercoiled DNA minicircles are templates which carry a single promoter. Such minicircles serve as templates to study the steps of transcription initiation from a unitary promoter and to synthesize both full length transcripts and aborted transcripts under physiological conditions in the absence of other contaminating promoter activities. The invention is directed to *in vitro* transcription of promoters, such as the *E. coli gal* promoter, using such minicircle DNA templates. The full-length transcripts from *gal* P1 and P2 promoters responded to cAMP-CRP in a manner identical to that observed *in vivo*. In addition, an *in vitro* transcription assay, using minicircle DNA containing the *gal* promoter with *lac* operators, was employed to elucidate the molecular mechanism of repression. Lac repressors, which associate into a tetramer and form a DNA loop, repressed transcription from P1 and P2, while a non-looping Lac repressor mutant failed to show normal repression of both of the *gal* promoters.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 5 C12N15/10 C12N15/70 C12N1/21 //(C12N1/21,C12R1:19)

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>BIO/TECHNOLOGY vol. 2, no. 12, December 1984, NATURE AMERICA, INC., NEW YORK, US; pages 1045 - 1049 K. BACKMAN ET AL. 'Use of synchronous site-specific recombination in vivo to regulate gene expression' see page 1045, right column, line 17 - page 1047, right column, line 14; figure 2 see page 1048, right column, line 4 - line 15</p> <p style="text-align: center;">--- -/--</p>	1-11

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*A\* document member of the same patent family

Date of the actual completion of the international search

21 April 1994

Date of mailing of the international search report

29 -04- 1994

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>GENE vol. 40, no. 1, 1985, ELSEVIER PUBLISHERS, N.Y., U.S.; pages 163 - 168 A.J. PODHAJSKA ET AL. 'Control of cloned gene expression by promoter inversion in vivo: construction of the heat-pulse-activated att-nutL-p-att-N module' see page 164, left column, line 23 - line 30 see page 164, left column, line 36 - line 39 see page 164, right column, line 31 - line 34</p> <p>----</p>	1-11
Y	<p>PROC. NATL. ACAD SCI. vol. 85, no. 24, December 1988, NATL. ACAD SCI., WASHINGTON, DC, US; pages 9683 - 9687 R. HABER AND S. ADHYA 'Interaction of spatially separated protein-DNA complexes for control of gene expression: Operator conversions' cited in the application see page 9683, right column, line 13 - page 9686, right column, line 26</p> <p>----</p>	1-8
A	<p>MOLECULAR BIOLOGY OF THE GENE, FOURTH EDITION 1987, BENJAMIN/CUMMINGS PUBL. COMP. INC., MENLO PARK, CA, US; WATSON, HOPKINS, ROBERTS STEITZ, WEINER see page 328, paragraph 2 - page 329, paragraph 2; figures 11-21, 11-24</p> <p>----</p>	1-11
A	<p>GENE vol. 56, no. 1, 1987, ELSEVIER PUBLISHERS, N.Y., U.S.; pages 145 - 151 N. HASAN AND W. SZYBALSKI 'Control of cloned gene expression by promoter inversion in vivo; construction of improved vectors with a multiple cloning site and the ptac promoter' see page 149, right column, line 15 - page 151, left column, line 5</p> <p>----</p> <p style="text-align: center;">-/--</p>	1-11

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	PROC. NATL. ACAD SCI. vol. 89, no. 23 , 1 December 1992 , NATL. ACAD SCI., WASHINGTON, DC, US; pages 11264 - 11268 H.E. CHOY AND S. ADHYA 'Control of gal transcription through DNA looping: Inhibition of the initial transcription complex' see page 11265, left column, line 24 - page 11267, right column, line 25; figure 1 ---	1-18
P,X	PROC. NATL. ACAD SCI. vol. 90, no. 2 , 15 January 1993 , NATL. ACAD SCI., WASHINGTON, DC, US; pages 472 - 476 H.E. CHOY AND S. ADHYA 'RNA polymerase idling and clearance in gal promoters: Use of supercoiled minicircle DNA template made in vivo' see page 473, left column, line 71 - page 476, right column, line 8; figure 1 -----	1-18